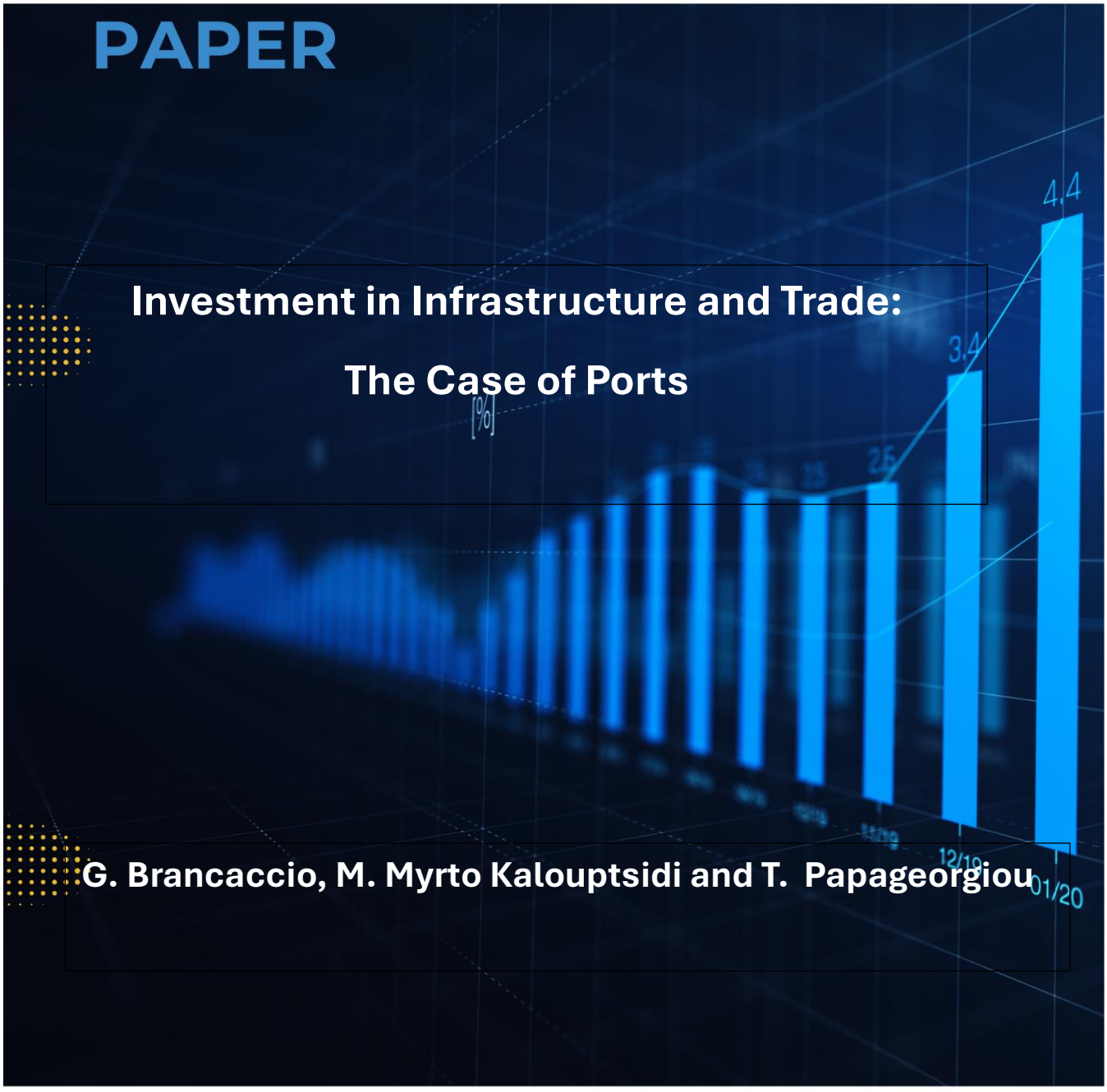
International Economics
& Development Laboratory

WORKING
PAPER

No. 12 | November 2025



Investment in Infrastructure and Trade:

The Case of Ports

G. Brancaccio, M. Myrto Kalouptsidi and T. Papageorgiou

HELLENIC REPUBLIC
National and Kapodistrian
University of Athens
DEPARTMENT OF ECONOMICS

HELLENIC REPUBLIC
National and Kapodistrian
University of Athens
DEPARTMENT OF ECONOMICS

Investment in Infrastructure and Trade: The Case of Ports

Giulia Brancaccio, New York University,

USA, giulia.brancaccio@nyu.edu

Myrto Kalouptsidi, Harvard University,

USA, myrto@g.harvard.edu

Theodore Papageorgiou, Boston College,

USA, theodore.papageorgiou@bc.edu

Working Paper

International Economics and Development Laboratory (IEDL),
University of Athens, Greece

November 6, 2025

Investment in Infrastructure and Trade: The Case of Ports

Giulia Brancaccio[*], Myrto Kalouptsidi[†], Theodore Papageorgiou[‡§]

New York University, Harvard University, Boston College

October 2025

Transportation infrastructure is vital for the smooth functioning of international trade. Ports are a crucial gateway to this system: with more than 80% of trade carried by ships, they shape trade costs, and it is critical that they operate efficiently. Yet ports are susceptible to disruptions, causing costly delays. With enormous budgets spent on infrastructure to alleviate these costs, a key policy question emerges: in a world with high volatility, what are the returns to investing in infrastructure? To address this question, we introduce an empirical framework that combines insights from queueing theory to capture port technology, with tools from demand estimation. We use our framework, together with a collection of novel datasets, to quantify the costs of disruptions and evaluate transportation infrastructure investment. Our analysis unveils four policy-relevant messages: (i) investing in port infrastructure can lead to substantial trade and welfare gains, but only if targeted properly– in fact, net of costs, the marginal return to investment is positive at a minority of US ports; (ii) there are sizable spillovers across ports, as investing in one port can decongest a wider set of ports, suggesting that decision-making should not be decentralized to local authorities; (iii) the economies of scale arising from queuing would lead a planner to concentrate investment in large, geographically dispersed megaports; (iv) macroeconomic volatility can drastically change returns to investment.

Keywords: transportation, infrastructure, ports, congestion, macroeconomic volatility, disruptions, spillovers

[*]New York University, Stern School of Business, Kaufman Management Center 44 West 4th Street New York NY 10012, USA, giulia.brancaccio@nyu.edu

[†]Harvard University, Department of Economics, Littauer Center 124, Cambridge MA 02138, USA, myrto@g.harvard.edu

[‡]Boston College, Department of Economics, Maloney Hall 332, Chestnut Hill, MA, 02467, USA, theodore.papageorgiou@bc.edu

[§]We are thankful to George Alessandria, Jim Anderson, Costas Arkolakis, Nick Buchholz, Arnaud Costinot, Dave Donaldson, Ed Glaeser, Paul Grieco, Gabriel Kreindler, Isabela Manelici, Marc Melitz, Kostas Metaxoglou, Amedeo Odoni, Ariel Pakes, Steve Redding, Marta Santamaria, and Edouard Schaal for their many helpful comments. We are extremely grateful to Constanza Abuin, Pedro Degiovanni, Toren Fronsdal, Peleg Samuels, Zachary Saunders, and Yixin Zhou for excellent research assistance. We also gratefully acknowledge financial assistance from the NBER Transportation Economics in the 21st Century Grant, as well as NSF Award SES-231533-231534-231535.

1 Introduction

Transportation infrastructure is crucial for the smooth functioning of international trade. As such, enormous funds are spent every year to improve the efficiency and resilience of the infrastructure network. Ports, which are the focus of this paper, are a vital link of this network. With more than 80% of trade (amounting to $11 billion tons and about $20 trillion) carried by ships, they are the literal gateway to international trade. Ports are therefore crucial determinants of trade costs, as became painfully apparent during the COVID-19 pandemic and its aftermath, when the disruptions to port operations upended supply chains, and generated enormous costs to importers and exporters worldwide, who faced massive delays in obtaining their goods.

Investment in transportation infrastructure has the potential to alleviate these costs. A key policy question therefore is: in a world with high volatility, what are the returns to investing in infrastructure? To answer this question, we first assemble a collection of novel micro-data on various aspects of port operations. We then introduce an empirical framework that combines insights from queueing theory to capture port technology, with tools from demand estimation. We use this framework to quantify the cost of disruptions and to evaluate transportation infrastructure investment. While our analysis focuses on maritime ports, in principle this framework can be applied to other setups with congestion, such as airports.

Our analysis unveils four policy-relevant messages: (i) investing in port infrastructure can lead to substantial trade and welfare gains, but only if targeted properly– in fact, net of costs, the marginal return to investment is positive at a minority of US ports; (ii) there are sizable spillovers across ports, as investing in one port can decongest a wider set of ports, suggesting that decision-making should not be decentralized to local authorities; (iii) the economies of scale arising from queuing would lead a planner to concentrate investment in large, geographically dispersed megaports; (iv) macroeconomic volatility can drastically change returns to investment.

More specifically, we leverage detailed micro-data on various aspects of port operations, obtained from a number of different sources. Using the universe of port calls, we collect data on port "outputs", which in our context includes port service times and queues, as well as on port demand. In addition, using satellite images, we collect information on port "inputs", namely existing port infrastructure such as berths, storage space and cranes, as well as port employment from union membership. We also obtain access to micro-data on the prices ports charge for their services. Finally, we collect information on the

cost of port expansion, such as land values and dredging costs, which are used by authorities to make investment decisions. Our study focuses on US ports and on dry bulk carriers, which carry about half of the volume of maritime trade. Dry bulk ships are the main mode of transportation for commodities, such as grain, ore, and coal.

Our key measure of port efficiency is the total time a ship spends at port, which includes both its service time, i.e. the time it takes to discharge or load cargo, but also the time it spends outside the port waiting to be serviced. Ships spend on average 109 hours per trip at US ports. This amounts to roughly half of the average travel time, suggesting that iceberg transport costs, that are a function of distance, miss a key component of trade costs. Time at port can be very costly, both because of the user cost of ships, but also because of any indirect costs of delays along the entire supply chain. As we describe below, our empirical model allows us to measure this cost of port disruptions.

In order to evaluate the role of transportation infrastructure we need two key ingredients: first, we need to accurately capture infrastructure technology, which we do through a port production function. This helps us understand both the dynamics of congestion, as well as how investment in infrastructure affects congestion. Second, we need a demand system for transportation services. This is necessary both because demand endogenously responds to infrastructure expansions, but also in order to quantify their welfare gains. We briefly discuss each ingredient in turn, which along with the cost estimates of infrastructure investment, allow us to perform our welfare analysis.

Using insights from queueing theory, we propose a micro-model of port technology that takes port inputs, such as infrastructure, labor, and cranes, and produces output, namely time at port. Our setup endogenously generates convex costs of congestion. To see this, note that at low levels of demand, time at port is almost flat, reflecting mostly service time; as demand keeps rising, time at port starts increasing and after a certain point, port infrastructure can barely keep up. Now even small increases in demand can make the queue explode, making ports susceptible to disruptions. Investment in infrastructure, by smoothing these fluctuations, can generate substantial welfare gains. Finally, the queuing technology leads to economies of scale in port size, so that all else equal, larger ports have lower wait time. Indeed, given the randomness in ship arrivals, it is possible that a port is congested, while a neighboring port has spare capacity; thus merging the two ports leads to lower expected wait time.

We then estimate a demand system for port services: potential exporter-importer pairs decide which port to use and value lower time at port, shorter distance to their origin and destination, lower port prices, as well as other characteristics of the port. Importantly, we allow for aggregate macroeconomic

conditions to shift demand over time. One complication in our analysis is that the ship movement data that we use does not contain information on the final US destination/origin of a good. For example, we may observe a ship load at the port of Houston (TX) and head to Europe, but we do not observe the origin of the shipment within the US. To address this, we augment our analysis with aggregate trade data on freight movements within the US. We estimate the demand system following the IO literature on demand estimation (Berry, Levinsohn, and Pakes, 1995, Berry et al., 2004).

Our estimates suggest that agents are quite sensitive to changes in time at port. Moreover we find that the average ship is willing to pay $64,000 to reduce the time at port by one day. This amount is substantially higher than the direct user cost of the ship, thus capturing the cost of port disruptions as they propagate through the supply chain. Although the geographic and other characteristics of the port do affect demand, we find that half the shippers do not choose the most geographically convenient port. This suggests that time at port acts as a barrier to trade, forcing shippers to sacrifice geographic convenience in order to avoid port congestion. Finally, substitution patterns vary greatly by port: for instance, ports in the Great Lakes substitute only locally, while ports in the Gulf and East Coast boast a wide range of substitutes as they serve a much wider set of origins and destinations.

Having all the necessary ingredients, we combine our estimates of technology with our estimates for port demand, as well as the costs, to perform a comprehensive cost-benefit analysis of investment in port infrastructure. We compute (i) gross (welfare gains) and net (welfare gains minus infrastructure costs) investment returns from capacity expansions at individual ports; (ii) how these returns shift in more volatile environments; (iii) the outcomes under a planner that invests sequentially in multiple ports. We discuss briefly each in turn.

First, we consider the following question: if we could increase the capacity of the US port system by one slot, i.e. a single port can now handle one additional ship, where would we allocate this additional slot? To put this in context, this increase corresponds to approximately a 1% increase of the total US capacity.

We find that this increase in effective port capacity increases US trade and exporters' and importers' welfare by as much as 1% and 0.8% respectively, but that the effect is quite heterogeneous across ports. The highest welfare gains are obtained for ports in the East Coast and the Gulf, such as Hampton Roads (VA) and Destrehan (LA). At the other extreme, investing in the Great Lakes barely makes a dent in total trade and aggregate welfare. This geographical dispersion in the returns from investment showcases that targeting can increase the gains from investment manyfold.

What makes a port generate high (vs. low) welfare gains? When capacity increases at a port, demand there spikes from an inflow of both new shippers, as well as those that substitute away from other ports.[1] The more geographically attractive, and the more congested the port, the higher the increase in its demand. For instance, Hampton Roads (VA), one of the ports with the highest returns, is both congested and able to serve many and diverse origins and destinations, which lowers the transportation costs that shippers need to shoulder.

But that is not all. In this inter-connected system, changing capacity in one port does not happen in a vacuum. Expanding capacity at a port reduces demand at substitute ports, which now become decongested. In other words, investment at one port has positive spillovers to other ports, which in our setup explain up to 23% of the final aggregate increase in trade for the ports with the highest gains. Welfare gains thus do not just depend on the attributes of the treated port, but also on which other ports it decongests. Hampton Roads (VA) decongests a large set of ports that are themselves geographically attractive and congested, amplifying its welfare gains. The presence of these spillovers raises important policy questions regarding the coordination of investment activities and suggests that unless decisions are made centrally, they are unlikely to internalize the spillovers across ports.

Moving beyond gross returns, where is port investment justified? Comparing the estimated welfare gains to the infrastructure costs, we find that *net* returns are positive for 22 out of 51 US ports, and are highest for the ports in the Gulf, such as Destrehan (LA) and Convent (LA), as well as for Hampton Roads (VA).

Next, can increasing macroeconomic volatility warrant more investment? Recent work has documented that extreme shocks are becoming more frequent (see Bloom, 2009, as well as the literature on the recent supply chain disruptions). An increase in volatility makes both very large and very low demand realizations more likely. While low demand realizations matter little in our context, high demand shocks can lead to sharp increases in congestion. As a result, higher volatility makes infrastructure investment more valuable: on average, welfare gains increase by 75% when volatility doubles, and net returns turn positive for 11 additional ports. At one extreme, welfare gains more than double for some ports. At the other extreme, for a third of ports net returns do not turn positive even when volatility more than triples; for these ports, investment is unlikely to be warranted even if the policymaker only cares about resilience.

Finally, we consider a planner that invests in multiple ports sequentially, as long as net returns remain positive. We find that the planner concentrates a large share of their investment in the three ports with

[1]The endogenous response of demand dampens the original fall in congestion. This is reminiscent of the "Iron Law of Congestion" (Downs, 1962); the increase in trade however, generates positive welfare gains.

the highest individual returns- Kalama (WA), Destrehan (LA) and Hampton Roads (VA). Overall, such an investment program would generate $8 billion in total net welfare gains in the form of increased trade through U.S. ports, with a total budget of $1.5 billion. Strikingly, the planner disperses investments across the US to minimize travel distances; however, at the local level, the planner creates a regional megaport. This harnesses the economies of scale arising from queueing: pooling berths together allows the planner to avoid wasteful situations where in one port berths are idle, while the other port has a queue of waiting ships, thereby reducing wait times.

Related Literature This paper relates to a broad and diverse literature studying the role of transportation and infrastructure on trade; this includes work in international trade, port development, urban transport and congestion, as well as supply chain disruptions.

A recent and influential literature has focused on the impact of transportation infrastructure on trade (e.g. Redding and Turner, 2015, Donaldson and Hornbeck, 2016, Redding, 2016, Donaldson, 2012, Fajgelbaum and Schaal, 2020, Allen and Arkolakis, 2022, Redding, 2025), as well as the evaluation of investment in infrastructure (Glaeser and Poterba 2021a and 2021b). Similarly to some of this work, we seek to quantify the welfare gains from investment in infrastructure, focusing on ports (instead of roads or railways). We contribute to this literature by explicitly modeling the dynamics of congestion and considering the role of macroeconomic volatility, while relying on a more granular methodology that captures the micro-structure of the infrastructure system.

A shorter literature has focused specifically on the impact of ports. Bonadio (2024) explores joint investment for roads and ports in the case of India. Brooks, Gendron-Carrier, and Rua (2021) and Ducruet, Juhasz, Nagy, and Steinwender (2024) focus on the historical context of containerization to study the impact of port development on city population growth. Bailey (2021) considers competition between US ports focusing on business stealing. Sytsma and Wilson (2020) consider the port choice of US importers. Fuchs and Wong (2025) explore investments in multi-modal transport networks relying on the methodology of Allen and Arkolakis (2022). See also Friedt and Wilson (2020) for a recent survey of the impact of port investment both on local economic activity, but also on the economy more broadly. Finally, a literature in port management and operations has studied related issues (Notteboom, Pallis, and Rodrigue, 2022). We extend the literature by performing a comprehensive cost-benefit analysis, while incorporating a queueing model for congestion and allowing for positive spillovers across ports and macroeconomic volatility.

We also contribute to a literature studying maritime transportation, e.g. Hummels and Skiba (2004), Hummels (2007), Hummels, Lugovskyy, and Skiba (2009), Kalouptsidi (2014), Ishikawa and Tarui (2018), Asturias (2020), Brancaccio, Kalouptsidi, and Papageorgiou (2020), Wong (2022), Brancaccio, Kalouptsidi, Papageorgiou, and Rosaia (2023), Ganapati, Wong, and Ziv (2024), and Heiland, Moxnes, Ulltveit-Moe, and Zi (2025).

In addition, we relate to a large literature that studies congestion, primarily in the context of urban transportation. In classic work, Vickrey (1969) and Arnott et al. (1990) provide models of road congestion, while more recently, a growing body of work embeds different types of congestion models in structural models of urban transportation (e.g. Durrmeyer and Martinez, 2022, Bordeu, 2023, Cook and Li, 2023, Conwell, 2024, Almagro, Barbieri, Castillo, Hickok, and Salz, 2024, Kreindler, 2024, Björkegren et al., 2025, Hierons, 2025). We rely on canonical models of queueing theory (see e.g. Larson and Odoni, 2007 and Bhat, 2008) and embed them in an empirical framework that allows us to analyze infrastructure investment. This provides a microfoundation for the congestion costs encountered in the literature above.

We also relate to a nascent but growing literature on supply chain disruptions (Carvalho, Nirei, Saito, and Tahbaz-Salehi, 2021, Grossman, Helpman, and Lhuillier, 2021, Elliott et al., 2022, Alessandria, Khan, Khederlarian, Mix, and Ruhl, 2023, and Bai, Fernández-Villaverde, Li, and Zanetti, 2024). Finally, this paper naturally relates to the large literature in international trade studying the role of geography in international trade, as well as the importance of trade costs in explaining trade flows between countries (e.g. Eaton and Kortum, 2002, Anderson and Van Wincoop, 2003, Melitz, 2003).

The rest of the paper is structured as follows: Section 2 briefly describes the setting, data, and documents facts regarding time at port. Section 3 introduces our framework for port technology and port demand. Section 4 presents the estimation procedure and results. Section 5 contains our welfare analysis, while Section 6 concludes.[2]

2 Setting, Data Description and Main Empirical Patterns

In Section 2.1 we describe the setting; in Section 2.2 we present the data sets used and some basic summary statistics; and in Section 2.3 we present facts about time at port.

[2]Appendix A provides additional figures and tables, Appendices B and C detail the construction of the infrastructure and cost data respectively, while Appendix D outlines the algorithm to compute the equilibrium of the model.

2.1 Setting

A port is a maritime facility, consisting of both water and land, where ships can load or discharge cargo. Upon arrival at the port, a vessel may have to wait in anchorage, a designated area outside the port. Once berth space is freed up, the ship enters the port to get serviced: its cargo is loaded or discharged by port workers and port capital equipment, such as cranes. The cargo is stored in the port's designated storage areas to be later moved on to trucks or rail to its final destination, and the ship departs.

A port consists of a "landlord" (usually the local port authority), who owns the land and part (or all) of the basic port infrastructure, which may include berths, (un)loading equipment, storage areas, buildings, etc. Sometimes, the landlord leases land to terminal operators who are usually private enterprises. These are responsible for the numerous port services a ship receives while at port, such as pilotage, cargo-handling, towage, etc. A ship receives on average 40 services per visit, for which it pays different charges. Although port charges can differ substantially across ports, they do not vary much over time, as discussed below.[3]

Investment in port infrastructure is the focus of heated policy debates. Between 2012-2016, $46 billion were dedicated to capital investments (of which $18 billion was public), while between 2021-2025, $163 billion expenditures were scheduled according to the American Association of Port Authorities (of which, the Infrastructure Investments and Jobs Act (IIJA) enacted in 2021 was around $17 billion). Port infrastructure involves water investments, which amount largely to dredging, as well as land investments, which relate to storage expansion.

In the US, port investment is usually funded by a mix of private and local public funds (e.g. through municipal bonds).[4] Investment decisions are undertaken by port (or other local) authorities and terminal operators who may have a mix of incentives including output or revenue considerations, job creation, local spillovers, and profit maximization. Port authorities face different oversight by state authorities; for instance, Oregon ports face substantial constraints by the state, as opposed to Florida ports where port authorities act largely independently.[5] Overall, in the US, ports should be thought of as quasi-public authorities, primarily aiming to serve local interests, such as high throughput, employment and connectivity to the area.[6]

[3] There are a few explanations for this, related to political incentives to facilitate trade and maintain stability and transparency, as well as to the complexity of pricing, which involves a long list of services.

[4] Waterways are an exception, as they are federally funded.

[5] https://www.oregon.gov/biz/Publications/Ports/2010PortPlan.pdf

[6] Several policies affect maritime trade. One of the most prominent ones, the Jones Act, restricts shipping between US ports to ships that are U.S.-built, owned, and crewed. Since we study international rather than domestic trade, this policy does not affect our analysis.

In this paper, we focus on commodities and raw materials, such as iron ore, steel, coal and grain, which are carried in dry bulk carriers. These are vessels designed to carry a homogeneous unpacked dry cargo, for individual shippers on non-scheduled routes. Bulk carriers operate like taxi cabs: each ship transports only one cargo at a time, for a trip between a single origin and a single destination.[7] The US is an important exporter of these goods and its ports handle a substantial amount of worldwide trade, especially for grain, which is exported worldwide from the ports in the Gulf. Overall, dry bulk shipping accounts for about half of total seaborne trade in tons (UNCTAD) and 45% of the total world fleet (but less in terms of trade value), which includes also containerships and oil tankers. The industry is unconcentrated, consisting of a large number of small shipowning firms (see Brancaccio, Kalouptsidi, and Papageorgiou (2020) for a detailed description of the industry, as well as a conceptual framework for its operation).

Last but not least, our setting is characterized by severe volatility: historically, shipping has been subject to notorious fluctuations. For instance, even though the disruptions experienced during the COVID-19 Pandemic were an extreme event, they was not incomparable to other recent disruptions. For instance, in the case of bulk carriers studied here, they were in fact mellower than the dramatic boom and bust of 2006-2008.[8] The disruptive nature of shipping is due to the combination of a demand (for transportation) that is subject to frequent local and global macroeconomic shocks, with a supply, driven by ships and ports, that is highly inelastic (see Brancaccio et al., 2025a, 2025b). The volatile nature of this sector is a crucial determinant of investment returns, and thus a key feature of our analysis.

Our analysis focuses on the largest 51 US ports by bulk throughput, covered by the datasets described below.[9] Ports are very diverse along several dimensions. Geography is a key determinant of a port's performance: its location and proximity to producing and consuming regions, but also whether it is in deep vs. shallow waters, inland vs. sea, its tide range, etc. The left panel of Figure 1, depicts the annual throughput of US ports and illustrates the extent of size heterogeneity: the smallest ports in this group

[7] It is worth noting that bulk ships are very different from containerships, which operate like buses: containerships carry cargos (mostly manufactured goods) from many different cargo owners in container boxes, along fixed itineraries/routes according to a timetable. It is technologically impossible to substitute bulk with container shipping. We choose to focus on bulk carriers because they offer a simpler setup to study port congestion (as would oil tankers). Indeed, container shipping features many complications such as hub-and-spoke networks of ports, vertical integration with port operators, market power against ports and others. Nonetheless the basic economic mechanisms discussed in this paper are useful for containers as well, as well as other settings with congestion, such as airports.

[8] In the mid-2000s, a rapid growth of raw material imports (a major demand shock), particularly in China, led to sustained increases in freight rates and a seven-fold surge in the new ship backlog between 2003 and 2008. The Great Recession in 2008 led to an idling of the existing fleet, while another 70% of that fleet was still scheduled for delivery by 2012 (Kalouptsidi, 2014). At the time, this was considered one of the most dramatic shipping cycles in history.

[9] We observe port calls for all 61 US ports that handle bulk. However, for 10 of these ports (which collectively account for 5.4% of total ship arrivals) we were not able to obtain data on labor or infrastructure, so we drop them from our analysis.

handle about 10 ship per year, while the largest handle more than 600 ships per year.

2.2 Data

We collect a number of novel datasets that are necessary to estimate the three key ingredients of our analysis: (i) the port production function, (ii) port demand, and (iii) the cost of port infrastructure.

Data for Port Production Function In order to measure how long ships spend at the port, we employ the universe of port calls, between 2016 and 2021 for bulk carriers larger than 10,000 DWT. The dataset is obtained from AXS Marine (AXS Dry). Each observation corresponds to a unique vessel and port, and reports a timestamp for the vessel's arrival at the port's anchorage, a timestamp for the vessel's entry to and the exit from the port, as well as an indicator for whether the ship loaded or discharged, and an estimate of the commodity on board. The port calls are created by AXS Marine from AIS (Automatic Identification System) data that report exact vessel positions (latitude and longitude) every few minutes.

In addition, we construct a dataset of global port infrastructure.[10] Because ports seldom share information about their infrastructure, we use geolocation data and satellite imagery (provided by Google Maps and Google Earth Pro) to collect this information. Focusing on the US, we collect data for each of the ports in our sample, from 2010 through 2021, and manually encode the following measures of port infrastructure: miles of each berth, acres of storage space, and number of cranes. In Appendix B we provide the details on how this dataset was collected and encoded.

Finally, to obtain a measure of port employment, we use publicly available annual data on longshoremen union membership reported to the Department of Labor in the US. We match this to ports based on the zipcode of each union.[11]

Data for Port Demand We leverage the universe of port calls to construct the monthly number of arrivals at each of the ports in our dataset. Because this dataset does not contain information on the final destination (or origin) of a shipment within the US, but only the US port used (and the foreign location

[10] Although there are publicly available datasets that report some measures of port infrastructure we concluded that the reported measures did not accurately reflect a port's capacity. To give one example, a commonly reported capacity measure is the number of berths; berths however can be short or long. For example, berths at the port of Baltimore (MD) range from 0.12 miles to 1.0 miles, and clearly offer different capacity. In addition, berths are not the only relevant feature of a port's capacity: storage space is also important, since limited storage space implies a ship may need to wait until storage is emptied out (e.g. by trucks) to further discharge. Cranes are another important measure of port infrastructure.

[11] This dataset is not as granular as we would like, since membership is reported annually, and it only includes longshoremen that are fully registered (non-casual) union members. We corroborate our findings using detailed employment data available for four large US ports (Oakland, Los Angeles, Tacoma and Seattle) by the Pacific Maritime Association. Indeed, as shown in Table 4 in Appendix A, total man-hours from shift-level information is highly correlated with our annual data.

where the ship came from), we complement it with the Freight Analysis Framework (FAF) Regional Database of the Bureau of Transportation Statistics, which contains annual data on freight movements between foreign regions and US states and metropolitan areas, including information on the region of entry (or exit) as well as the product transported.

Additionally, we use a dataset of port charges, from a provider that brokers such transactions and who wishes to remain anonymous. This is a sample of port calls globally that covers approximately 35% of port calls, with the associated breakdown of the charges for each service requested (of which the main ones are berth dues, which include cargo handling dues, line handling, and pilotage). The dataset runs from 2016 to 2020. Although we cannot guarantee that this sample is not selected, we find that it has similar geographical and ship size coverage as our universal port call data.

Data for Cost of Port Infrastructure Port expansion amounts to constructing an additional section of berth to service incoming or outgoing freight. The cost of such an expansion consists of two main elements: (i) dredging; and (ii) acquisition of land for storage. For the former, we use publicly available cost projections from US Army Corp of Engineers (USACE) port projects. From these, we produce site-specific cost estimates for dredging a new berth, taking into account the rock vs. non-rock content of each port-berth's seafloor, obtained from the US Geological Service's (USGS) National Crustal Model. For the cost of land acquisition, we first use open-source map data from OpenStreetMap, to identify parcels of industrial and commercial land near each existing berth from our hand-collected geolocational data from Google Earth. We then use highly granular land value estimates from the ecology and conservation literature (Nolte, 2020) to assign unit land prices to these parcels.[12] We discuss our approach in detail in Section 4.3, as well as Appendix C.

Summary statistics Our final port call dataset involves 42,341 port calls between 2016 and 2021, for 7,499 ships and 51 ports. As shown in Table 1, on average, ports handle 164 port calls (and about 7.1 million tons) annually, though this can vary substantially (as shown also in the left panel of Figure 1). Ships spend on average 109 hours at port; of this, on average 67% is attributed to service time and 33% to wait time at anchorage (i.e. time spent in congestion). However, time at port varies dramatically across ports and years. Finally, it is also worth noting that wait time at anchorage is substantially more volatile than service time at the port level.

[12] A similar approach was taken by Ducruet, Juhasz, Nagy, and Steinwender (2024) who show the criticality of land scarcity and assumed the port expansion cost amounts to land costs. We extend their cost analysis by adding dredging costs, which amount to 32% of total costs based on our estimates.

	Mean	SD	Min	25th percentile	75th percentile	Max
Annual number of port calls	164	137	2	68	214	860
Annual throughput (tons)	7,116,448	6,278,038	17,670	2,754,136	9,062,871	33,138,942
Time at port (hours): average	109	59	8.4	67	145	334
Time at port (hours): st. dev. over time	85	54	2.8	46	114	400
Wait time (hours): average	33	42	0	0	53	273
Wait time (hours): st. dev. over time	52	56	0	0	90	280
Fraction of port calls with positive wait time	0.3	0.29	0	0.057	0.57	0.87
Port dues ($ per ton)	1.9	1.4	0.097	0.9	2.6	12
Port dues (in $) per port call	95,608	63,910	6,467	45,638	147,208	303,286
Length of berths (miles)	1	0.83	0	0.43	1.5	3.1
Storage space (acres)	183	172	4.8	68	234	814

Table 1: Summary statistics. Port calls (throughput in tons) is the total annual number of ships (total tonnage) that were handled at port. The average time at port refers to time at port averaged over time, while standard deviation of time at port is computed over time per port. The fraction of port calls with positive wait time is computed for each port and year.

Total port charges are on average $95,608 per port call. Port prices are heterogeneous over ports, but not over time. Indeed, in Table 3 in Appendix A we regress port prices on several observables (e.g. port, month, commodity, ship size) and demonstrate that port fixed effects soak much of the price variation, in contrast to time fixed effects that tend to explain little (similarly, there is no evidence of additional charges during congested periods).[13]

On average, ports have 1 mile of berths (sd 0.8) and 183 acres of storage (sd 172), however the heterogeneity in infrastructure mirrors that in size; e.g., the port of Charleston (SC) has 27 acres of storage and 0.14 miles of berths, while the port of Houston (TX), one of the largest in the US, has 290 acres of storage and 1.9 miles of berths.

Finally, ports do not tend to specialize in specific commodities; 77% of ports receive all major bulk commodities including grain, ore, and coal (note that containers are generally handled in different terminals).[14]

2.3 Time at Port

Our key metric of port efficiency is the total time a ship spends at a port, defined as the sum of the wait time, i.e. the time a ship spends waiting in anchorage and the service time, i.e. the time a ship spends inside the port discharging or loading its cargo. Time at port is naturally a key determinant of port throughput.

[13]We also find that prices are not correlated with any port characteristics such as capacity, congestion, geography, etc. In ongoing work we explore the impact of optimal congestion pricing.

[14]Using the granular data on port operations available for West Coast ports from PMA, we test whether bulk operations are independent of container operations within a port; see Table 6 in Appendix A. In particular, we find that the number of container ships serviced in a port in a given shift does not affect the service time for bulk ships serviced in the same shift.

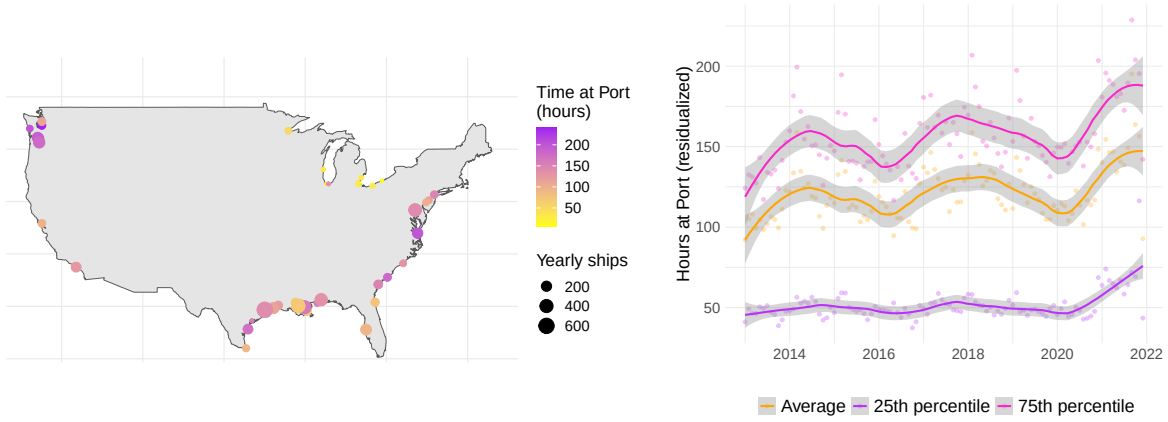


Figure 1: Port throughput and time at port over space and time. The left panel plots the average time at port in the US. It also indicates the size of the port in terms of annual throughput, designated by the size of its dot. The right panel plots the average time at port for each month, as well as its 25th and 75th percentile. We control for ship size and commodity on board.

Time at port is costly to ships and cargo-owners. The 109 hours spent on average at US ports constitutes 22% of travel time, suggesting that iceberg transport costs that are a function of distance miss a key component of trade costs (almost half -roughly 44%- if both loading and unloading ports are taken into account).[15] At the average shipping price of $14,000 per day during our sample, time at port has a direct cost of $68,250 per port call, plus any penalty fees ("demurrage") paid to/from shipowners in case of delay. Naturally, this is only a lower bound on the costs accrued, since it does not include delays for the next intermodal steps (usually trucks or rail), nor inventory costs that importers and exporters face. The demand estimates in Section 4.2 allow us to quantify directly the cost of these disruptions.

Although informative, the average time at port masks massive dispersion over both space and time. The left panel of Figure 1, depicts the average time at port in the US and demonstrates the sizable differences across ports: annual average time at port can range from 8 to 334 hours (see also Table 1). For the average (median) port 41% (46%) of the ships that arrive face a queue; however in some ports, such as Corpus Christi (TX), ships face queues more than 67% of the time, while in others, such as Cleveland (OH), they rarely wait in anchorage, if ever. This large heterogeneity is reflected in the different returns from investing in infrastructure, discussed in Section 5.

Similarly, time at port exhibits significant fluctuations over time, as shown in the right panel of Figure 1. Perhaps not surprisingly, the largest uptick in time at port occurs during the COVID-19 pandemic whereby ports experienced substantial increases in congestion. Indeed, the fraction of ships that had to wait almost doubled, going from 36% in 2019 to 77% in 2021. It is worth noting that these graphs do

[15]The correlation between time at port and the weighted distance of a port's trips is approximately 0.25, suggesting that distance is not a good proxy for port costs.

not capture the large rise in congestion observed in 2022 which is outside of our data sample. Finally, a key feature of the port network is that ports tend to be congested at the same time: as an example, for Baltimore (MD) and Hampton Roads (VA), two of the largest ports in the East Coast, the correlation of wait time is 0.48. This pattern suggests that macro volatility is crucial and generates correlated shocks across ports.

3 Framework

In this section we present our basic framework. Importer-exporter pairs decide whether to trade between a foreign and a US domestic location, and which US port to use. This choice depends on the geographical location of the ports, congestion there, as well as price and port quality. When the ship arrives at port, it either enters the anchorage and joins the queue, or it proceeds directly for service, if there is no queue. Upon completion of service, ships depart. The port queue evolves endogenously as ships arrive at port or depart after service. We first describe port technology, and then the demand that ports face.

3.1 Port Technology

Port technology transforms port inputs, such as labor, cranes and infrastructure, to port output, i.e., time at port. Consider a ship arriving at port. As already discussed, its time at the port is given by the sum of the service time, i.e, the time it takes for the ship to be loaded or discharged, and the wait time, i.e. any time the ship spends in anchorage, waiting to begin service.

We formalize this by adopting one of the most widely-used queueing models, the so-called the M/M/K model, where ship arrivals follow a Poisson process and port service times are exponentially distributed (see e.g. Larson and Odoni, 2007 and Bhat, 2008). Suppose that a ship arrives at port and finds Q ships ahead, either waiting in anchorage or at berth, each with expected service time T, while the port can handle at most K ships at a time. Then the ship's expected total time at port is equal to,

$$\underbrace{T}_{\text{service time}} + \mathbb{I}\{Q \geq K\} \underbrace{(Q - K + 1)\frac{T}{K}}_{\text{wait time (queueing)}} \tag{1}$$

In words, if the ship arrives and there are fewer than K ships at port (i.e. $Q < K$), then the port is operating under capacity, and the ship is able to be serviced immediately, so that total time at port simply equals the service time, T. If instead $Q \geq K$, then the ship needs to wait for a berth to open up

and the total time is augmented by the wait time.

To understand the expression for wait time in equation (1), note that if the port could handle only one ship at a time ($K = 1$), then the expected wait time would just equal QT, as the ship would need to wait for Q ships to complete service, each taking on average T periods. In the case where $K > 1$, the arriving ship does not have to wait for all Q ships ahead to complete service, since the port can handle more than one ship at a time. For example, if $Q = K$, then only one ship needs to complete service for a berth to open up, regardless of how large Q is. In general, $Q - K + 1$ ships need to conclude service before the ship can get to berth.[16] Since service times are exponentially distributed, a berth becomes available in T/K periods in expectation when K ships are being serviced. Therefore the expected wait time is $(Q - K + 1)(T/K)$.

Expected service time at port is given by

$$T(L/s, c/s, \omega) \tag{2}$$

and depends on the port's workers, L, cranes c, productivity ω, as well as the number $s \leq K$ of ships currently serviced. This specification is driven by the empirical finding that a ship's service time depends on how many other ships are being serviced (see Table 5 in Appendix A), suggesting that ports to some extent stretch their inputs across ships handled.[17]

In summary, time at port has three key components: service time T, port capacity, K, and the number of ships ahead, Q which is endogenous and depends on T, K, as well as the demand for port services. In our empirical implementation, we allow each of these objects to vary across time and ports.[18] From equations (1) and (2) the differential impact of different inputs starts to appear: infrastructure in port

[16] To see this, suppose that $K = 3$ and there are $Q = 7$ ships ahead, so that three ships are currently being serviced and there are four ships ahead in anchorage. Then, the three ships currently at berth need to conclude service, and be replaced by three of the ships in anchorage, leaving only one ship ahead in anchorage. At this point, a berth becomes available for the arriving ship, once two of the ships at berth conclude service. Therefore, in total five ships need to complete service before a berth becomes available.

[17] There are some implicit assumptions behind our formulations in (1) and (2), all of which have been verified empirically. In particular, we do not allow service time, T, to be affected by the number of ships in line (see Table 5 in Appendix A). Moreover we do not allow ships to skip the line so that the line is first-in-first-out (FIFO), as we find that in the US queue jumping occurs for less than 20% of observations. Finally, we do not allow port capacity K to adjust depending on the number of ships waiting.

[18] Throughout the model, we abstract away from ship size. Indeed, the number of "servers", K, would not be sufficient to describe the port's capacity, and one would have to consider the different combinations of ship sizes that could "fit" in the port at a single point in time. At the same time, each ship would need to keep track of the distribution of ships ahead of it, not only their number. This is a complex problem that even modern ports and airports do not necessarily perfectly optimize over. At the same time, capturing ship size does not impose first-order limitations to our analysis. Indeed, the focus of the paper is to compute welfare gains of expanding a port's capacity, holding fixed the distribution of ship sizes that a port handles.

capacity, K, matters for the port queue and therefore waiting time, whereas workers and cranes primarily affect the service time of each ship.

Finally, we discuss how the port queue evolves endogenously as ships arrive, are serviced, and leave. To this end, denote by λ the ship arrival rate and by $\mu(s) = 1/T(s)$, the rate at which a ship completes service, where we suppress the dependence of $T(s)$ on L, c, ω.[19] Then, the evolution of the queue follows a Markov structure. Indeed the probability $p_n(\tau)$ that in instant τ there are n total ships at port, either being serviced or waiting, satisfies,

$$\frac{dp_n(\tau)}{d\tau} = \begin{cases} -\lambda p_0(\tau) + \mu(1)\, p_1(\tau), & \text{if } n = 0 \\ -(\lambda + n\mu(n))\, p_n(\tau) + \lambda p_{n-1}(\tau) + (n+1)\,\mu(n+1) p_{n+1}(\tau), & \text{if } 0 < n < K \\ -(\lambda + K\mu(K))\, p_n(\tau) + \lambda p_{n-1}(\tau) + K\mu(K) p_{n+1}(\tau), & \text{if } n \geq K \end{cases} \tag{3}$$

In words, consider the case where $n < K$ (a similar argument holds for $n \geq K$): the change in the probability that there are n ships at port increases by (i) the probability that there are $n-1$ ships at port at τ, $p_{n-1}(\tau)$, and a new ship arrives, which happens with probability λ; and (ii) the probability that there are $n+1$ ships at τ, $p_{n+1}(\tau)$, and one of them is serviced, which happens with probability $(n+1)\,\mu(n+1)$. In addition the probability $p_n(\tau)$ decreases by the probability that there are n ships at port, $p_n(\tau)$, and a new ship arrives or one is serviced. Therefore, simulating the queue evolution boils down to solving this system of differential equations. We describe these simulations in detail in Appendix D.[20]

3.1.1 Understanding Congestion Dynamics

This framework sheds light on the formation and evolution of congestion, which we demonstrate through a series of simple simulation exercises. These insights will prove useful in understanding the impact of investment in infrastructure.

[19] We suppress the dependence of μ and T on L, c, ω, to ease exposition, as we assume that these variables change at a much lower frequency than the queue at the port, which changes with each ship arrival. We also suppress the dependence of the ship arrival rate, λ, on time, though in practice we allow demand to vary over time as we discuss in Section 3.2 (see also discussion in section 4.11 of Larson and Odoni, 2007). See Appendix D for details.

[20] We require that there is a high enough queue length $\bar{L}$, above which ships choose not to join the queue and $\lambda = 0$ in this case. Thus,

$$\frac{dp_{L+K}(\tau)}{d\tau} = -K\mu(K)\, p_{L+K}(\tau) + \lambda p_{L+K-1}(\tau)$$

In practice, we choose $\bar{L}$ to be very large so that it is never binding.

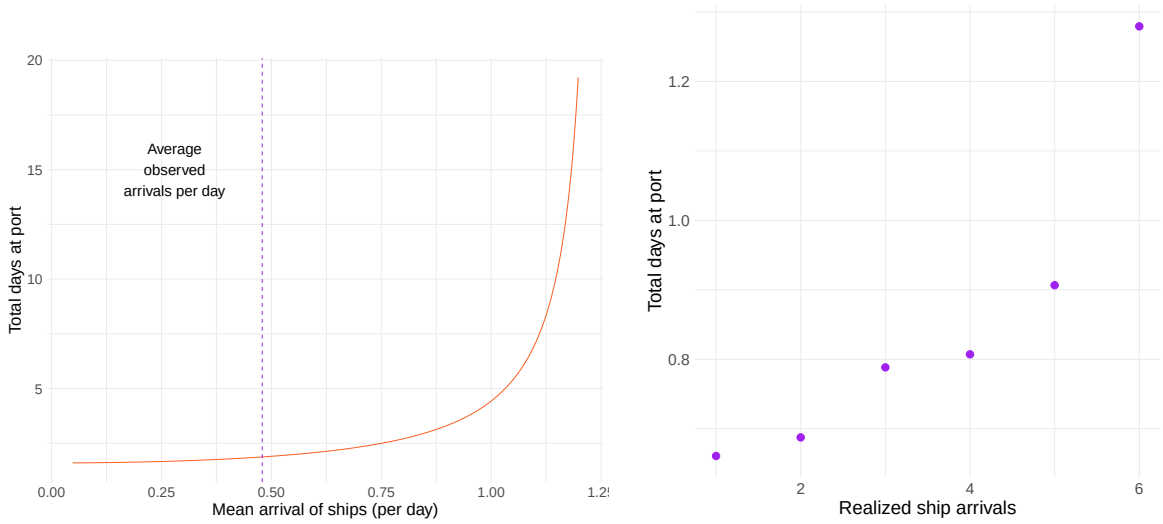


Figure 2: Convex cost of congestion. The left panel of this graph plots the relationship between time at port (vertical axis) and the mean arrival rate (horizontal axis) and is produced from model simulations of increasing the mean arrival rate of ships. In particular, for each level of the mean arrival rate on the horizontal axis, we simulate the port queue using the system of differential equations (3). We use the average service time, and the average capacity K. The right panel plots an example of this convex cost of congestion for one port in our data, the port of Duluth (MN). In particular, it plots the average time ships spend waiting (vertical axis) against a moving average of realized daily ship arrivals (horizontal axis).

Convex Congestion Cost Function First, consider the relationship between port congestion and demand. The left panel of Figure 2 plots the simulated time at port for the average port, for different values of the ship arrival rate.[21] Since time at port is costly, this graph captures the cost of congestion.

In periods with low average demand, time at port is fairly low: as demand is low, it is likely that the port is not at capacity, so an arriving ship immediately begins service and total time at port simply equals the service time. Note that increases in demand do not initially change time at port much, since the probability that the port is full remains low. After a certain point, however, it becomes more and more likely that an arriving ship has to wait at anchorage and total time at port increases, as the wait time becomes positive. Further increases in demand have large effects on wait time, as the port infrastructure becomes overwhelmed and the queue explodes.[22] In other words, even if we assume that for customers the cost of waiting at port is linear in time at port, the port production function endogenously generates a convex cost function for congestion.[23]

This convex congestion cost function explains why ports are susceptible to disruptions: at low levels

[21] We construct the average port as a port with average capacity, ship arrival rate, and service time. For each level of the mean arrival rate, we simulate the queue using equations (3).

[22] Formally the ratio of ships entering the system over the ships exiting approaches one, in which case the port queue approaches infinity.

[23] Lee and Cohen (1983) show formally that in the M/M/K queueing model, the expected wait time is convex in the arrival rate, λ. To our knowledge, no one has previously considered the economic implications of this finding (see discussion below).

of port utilization, demand shocks have little impact on the total time at port; on the contrary, when the port is already at capacity, further increases in demand can have large effects on the queue and ships' wait time. Therefore, an implication of this model is that even small demand shocks can lead to significant disruptions, if they occur when the port is already close to capacity. By the same logic, the accumulation of shocks can prove detrimental for port congestion.

This model formulation provides a microfoundation for the congestion costs encountered in the literature. For instance, in road congestion models, congestion often depends parametrically on the current demand for road usage. A notable difference is that here the cost of congestion depends not only on current demand for port services, but also on past demand realizations, allowing for endogenous congestion dynamics.

Although this analysis is based on our queuing model, we find that this convexity of congestion holds in the data as well. We depict one example in the right panel of Figure 2, which shows this relationship for the case of the port of Duluth (MN). The figure plots a moving average of daily ship arrivals on the horizontal axis and the average time ships spend at port on the vertical axis. As predicted by our framework and the simulations, in periods of low demand, ship wait time is zero and time at port remains low, as the port has spare capacity and ships immediately begin service; as demand increases, ships are more and more likely to have to wait. Interestingly, as demand increases further, wait time seems to explode, resulting in the expected convex relationship between time at port and port congestion.

The Role of Variance Next we examine the impact of an increase in demand volatility. Formally, we compute the average monthly arrival rate across ports and simulate wait times under a mean preserving spread for this process, where the dispersion of these monthly arrival rates increases. This implies that both very low and very high demand realizations now become more likely. Note that, due to the convexity of time at port with respect to ships' arrivals, high and low demand realizations have an asymmetric impact on time at port. On one hand, time at port is relatively flat for low realizations of ships' arrivals; on the other hand, high demand shocks dramatically increase congestion, overwhelming the system. As a result, increasing the variance of demand increases the average time at port that ships face. This is illustrated in the left panel of Figure 3 and suggests that it is not just the level of demand, but also its variance that increases the cost of congestion.

Convexity and Infrastructure How does investment in infrastructure impact congestion costs? Consider an expansion in infrastructure, i.e. an increase in port capacity K. As shown in the right panel

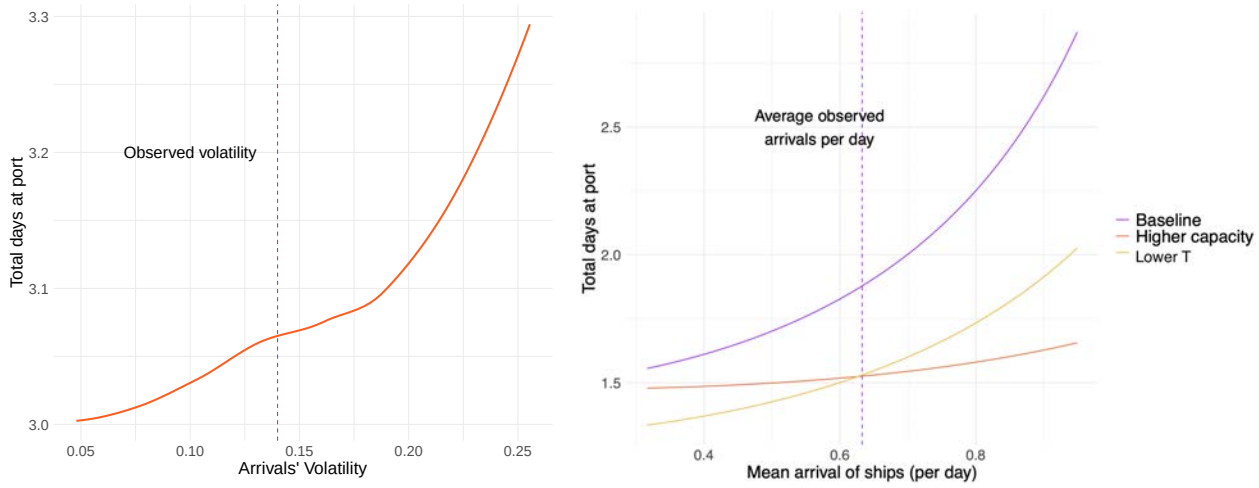


Figure 3: Congestion dynamics: the role of volatility and different port inputs. The left panel plots the relationship between time at port (vertical axis) and the variance of ships' arrival rate (horizontal axis). It is produced from model simulations of increasing mean preserving spreads of the arrival rate of ships. The right panel plots the relationship between time at port (vertical axis) and mean arrival rate (horizontal axis) when service time T falls or capacity K increases.

of Figure 3, the impact of this expansion is highly asymmetric for low vs. high demand. Indeed, higher capacity leads to very little change in time at port in states of low demand: time at port is roughly equal to the service time, which remains unaffected when K increases. When demand is high, on the other hand, there can be a substantial impact on time at port, as wait times fall. Much like a super market queue would disperse if a new cashier opens up, port wait time is slashed when new infrastructure comes online. Capacity, K, essentially only matters in periods where the port is already constrained, acting as a form of insurance against periods of high demand realizations, that can smooth out volatility.

It is worth contrasting the impact of infrastructure to that of a decline in service time T, which from (2) can be achieved by an increase in some combination of labor L, cranes, c and/or port productivity, ω. A reduction in service time is beneficial at all levels of demand, as shown in the right panel of Figure 3. At low levels of demand where total time at port is largely comprised of the service time, there is an immediate reduction in total time. At higher levels of demand, total time at port falls as well: beyond own service time being lower, the service time of ships ahead in the queue also falls, thus reducing wait time as well, leading to a substantial reduction in time at port. Hence, reductions in service time, generated from higher variable inputs, lead to symmetric reductions of congestion costs, in contrast to the case of expansions in infrastructure K.

Economies of Scale in Ports Finally, it's worth pointing out that the queuing technology leads to economies of scale in port size. In other words, all else equal, larger ports have lower wait time. This is

due to the combination of port capacity constraints and the lack of coordination among shippers, who face uncertainties in their travel.[24] To see this, consider two ports next to each other. Given the randomness in arrivals, it is possible that one port is congested, while the other has spare capacity; thus merging the two ports leads to lower expected wait time. Formally, we can show that doubling a port's capacity, K, and ship arrival rate, λ, while keeping expected service time, T, unchanged, leads to lower expected wait time. In fact, the expected wait time of the new port is less than half the wait time of the original port (as along as $\frac{\lambda T}{K} < 1$, i.e. the system is stable and the queue does not explode).

3.2 Demand for Ports

We next turn to demand for port services, which is necessary to estimate the welfare gains from port expansions.

There are D locations in the US, including both coastal and inland regions, and F locations abroad. In every foreign location f, there are M_f potential importers or exporters, each paired to a counterpart in location d. Each of them decides whether to trade a ship-sized shipment with d, and which port to use. In the case of exports, d is a US producing location and f is a foreign destination, while in the case of imports, d is a US importing destination and f is a foreign producing location.

Let t denote the month-year period. The payoff to shipment i, i.e., importer-exporter pair, traveling between f and d, from choosing port j in period t is given by,

$$u_{ijt} = \beta_F \text{dist}(f,j) + \beta_D \text{dist}(j,d) - \beta_T TT_{jt} - \beta_P p_j + \gamma_t + \gamma_f + \gamma_{l(j)} + \xi_{jft} + \epsilon_{ijt} \tag{4}$$

where $\text{dist}(f,j)$ denotes the distance between the foreign location f and port j and $\text{dist}(j,d)$ is analogously defined, TT_{jt} is the total time at port j in period t, and p_j is the price that port j charges for its services. The $\beta_.$ parameters capture how much the pair values each characteristic. Moreover, γ_t is a month fixed effect capturing macroeconomic fluctuations, while γ_f and $\gamma_{l(j)}$ are respectively fixed effects for the foreign location and the port's region (namely, either the Gulf, the East Coast, the West Coast, or the Great Lakes). Finally, ξ_{jft} is an unobserved demand shock for period t, that captures potentially time-varying unobserved port characteristics, such as its hinterland connectivity, reliability/management, as well as tariffs specific to foreign location f, while ϵ_{ijt} is an iid shipment-specific shock that follows a Type I extreme value distribution. If the shipper chooses not to trade to the US, they obtain a payoff that is

[24] Indeed there is substantial uncertainty in ship arrivals at port. A common saying in the industry is that "a ship is not a train" in that a ship cannot guarantee too far in advance the precise day it will arrive to a specific port. Indeed, weather alone substantially affects travel times.

normalized to zero, so that $u_{i0t} = 0 + \epsilon_{i0t}$.[25]

Note that TT_{jt} is endogenous since it depends on port demand, and therefore determined in equilibrium (defined in Section 3.3). In particular, it depends on the decisions of all other shippers; in equilibrium every shipper can correctly predict port flows and infer TT_{jt}.

The specification in (4) allows us to measure the welfare cost of delays, since the term $\beta_T TT_{jt}$, which captures the cost of port disruptions for the importer-exporter pair, also reflects the cost of delays along the entire supply chain. In addition, our specification allows for a rich demand response to an increase in time at port: a shipment's foreign and domestic locations shape how substitutable port j is when hit by disruptions. For example, when a shipment is headed to a final destination near port j, an exporter may be more willing to endure the long wait times rather than switch to a faraway port. Conversely, the shipper might be more willing to switch ports when the final destination is in the middle of the country.

Ultimately, the degree of substitutability across ports is an empirical question. In practice, demand cross-elasticities depend on whether shippers are able to substitute ports given ground transportation arrangements and establishment locations. As shown below, we find that the elasticity with respect to disruptions is substantial, while substitution has non-trivial geographic patterns, and this is a key driver of welfare gains.

Trade Flows The payoff specification (4) implies that the probability, σ_{jfdt}, that a pair trading between foreign location f and domestic location d chooses port j is,

$$\sigma_{jfdt} = \frac{\exp\left(\delta_{jft} + \beta_D \text{dist}(j,d)\right)}{1 + \sum_l \exp\left(\delta_{lft} + \beta_D \text{dist}(l,d)\right)} \tag{5}$$

where following Berry, Levinsohn, and Pakes (1995) it will prove useful to define,

$$\delta_{jft} \equiv \beta_F \text{dist}(f,j) - \beta_T TT_{jt} - \beta_P p_j + \gamma_t + \gamma_f + \gamma_{l(j)} + \xi_{jft} \tag{6}$$

which captures the "mean utility" of port j for shipments from/to f in month t. The probability of the outside option, σ_{0fdt}, is similarly defined. This implies that every month the total flow through j is given by $\sum_f \sum_d \sigma_{jfdt} m_{fd} M_f$, where m_{fd} is the fraction of potential importers-exporters in f paired with a counterpart in d, so that $\sum_d m_{fd} = 1$. Looking ahead at the estimation, one complication of this

[25]The outside good captures the option of not shipping at all between f and d, which would amount to trading the commodity with other counterparts, or not at all. It may also capture using other modes of transportation, or changing the timing of the shipment. We discuss this further in Section 3.4.

framework is that our data on ship flows, do not contain information on a shipment's US location d, so we do not directly observe the empirical analogue of σ_{jfdt}, i.e. $P(j|f,d,t)$; we discuss this further in Section 4.2.

Finally, the specific timing of the $\sum_f \sum_d \sigma_{jfdt} m_{fd} M_f$ ships that pass through port j within month t is random. In particular, a ship arrives at port j at Poisson rate $\lambda_{jt} = \sum_f \sum_d \sigma_{jfdt} m_{fd} M_f$; as such the demand model in this paper can be thought of as a micro-foundation of the Poisson arrival rate of ships at ports, which is an input to the queuing model capturing port congestion.

Macroeconomic Volatility Since ports are long-lived infrastructure assets, it is necessary to calculate the long-run returns from investment. This requires making demand forecasts. At the same time, as already discussed, volatility is a key feature of this setting. We thus enrich our framework in order to capture the evolution of aggregate demand. In our model, macroeconomic conditions are captured by the term γ_t. This term changes over time and is common across all ports j; intuitively, therefore, γ_t determines the choice of whether to export or not.[26] To capture the evolution of these macro shocks we assume that

$$\gamma_t = \rho_0 + \rho_1 \gamma_{t-1} + \epsilon_t, \tag{7}$$

with $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. In Section 4.1 we describe how we augment our data to estimate the times-series process for γ_t.

3.3 Model timing and equilibrium

In this model, each period (a month) is characterized by the aggregate conditions γ_t, and for each port, an expected time at port, service time, unobservable ξ_{jft}, price, as well as port-specific shocks, ϵ_{ijt}. The timing of the model is as follows:

1. At the beginning of period t, M_f potential importer-exporter pairs decide whether to trade and which port to use, taking into account the expected time at port TT_{jt} for all ports j. This determines the rate at which ships arrive at port j, $\lambda_{jt} = \sum_f \sum_d \sigma_{jfdt} m_{fd} M_f$.

2. Within the month, the queue evolves (at half-hour increments) as each of the ships arrives at port j. Equations (3) describe formally this Markov process. The resulting queue determines the expected time at port in month t for each port j.

[26] In fact, it would have been equivalent to define the payoff of the outside option as $u_{i0t} = \gamma_t + \epsilon_{i0t}$. Here, we subtract γ_t from the utility of all alternative options j, and set the mean payoff from the outside option equal to zero, so that $u_{i0t} = \epsilon_{i0t}$.

3. At the end of the month, macroeconomic conditions evolve according to the AR(1) process (7). Moreover, the shocks ξ_{jft}, and ϵ_{ijt} are drawn.

In each period t, an equilibrium consists of port market shares, σ_{jfdt} for each pair f, d, given by (5), and time at port TT_{jt} for all ports, such that (i) ships' expectations for time at port coincide with TT_{jt}; (ii) TT_{jt} is determined by the Markov process in (3) using the equilibrium arrival rates, λ_{jt}, implied by the equilibrium market shares, σ_{fdjt}, for all j.

3.4 Discussion

We close this section with a discussion on several of our assumptions and some caveats.

First, our model does not allow shippers to adjust the timing of their shipments. Yet this could be an important feature of reality, if **intertemporal substitution** occurs frequently due to expectations about future congestion, inventory management, etc. In our model this option is subsumed in the outside good. Unfortunately, our data does not allow us to estimate the extent of intertemporal substitution: our analysis is based on shipment-level data, while this would require information on shipping timing decisions made by the exporter-importer pair.

That said, we have some reasons to think that abstracting away from the timing of shipments is not a first order issue. First, we perform a simple test: we check whether demand (ship arrivals) in period t increases with congestion in the past, so that if a port is congested, shippers postpone shipment.[27] We find that there is no significant impact, across several specifications. Second, empirical evidence suggests that inventory costs are substantial, making delaying shipment potentially unattractive. According to Alessandria et al. (2010) "in the US, inventory costs on average account for 6 percent of sales, more than half of the 10 percent accounted for by labor".

In addition, our model does not consider **general equilibrium effects**. In other words, we do not allow the prices of commodities, nor other factor prices, to adjust after port investment. In our model, these effects are only indirectly captured by the elasticity with respect to time at port, β_T, which however is fixed in counterfactuals. This is an interesting avenue for future research, which we explore in Brancaccio et al. (2025b). Similarly, we do not allow the spatial distribution of economic activity across countries to change by port investment. In a similar vain, our counterfactuals do not let port prices, productivities,

[27] In particular, we use the specification $\lambda_{jt} = \beta_1 TT_{jt} + \beta_2 TT_{jt-1} + \epsilon_t$, where λ_{jt} are ship arrivals at port j in month t and TT_{jt} is time at port at j at t, and test whether β_2 is significant, instrumenting for both TT_{jt} and TT_{jt-1} as below. If intertemporal substitution is important, we should obtain $\beta_2 > 0$ and significant: in other words, if a port was congested then we might expect shippers to postpone their shipments to the next period.

or the internal transportation network to adjust after investment.

Finally, the **outside good** here may capture a few different choices that exporter-importer pairs have when not shipping between f and d. First, it includes the option of not shipping at all between f and d, which would amount to, for instance, selling the cargo domestically. Second, it captures other modes of transportation. However, for most bilateral trading regions (f, d) there is no alternative to oceanic shipping (these goods are not transported by air); in addition, we allow different foreign locations to have different access to alternative transportation modes by including fixed effects for the foreign location f. Finally, the outside good includes the option of waiting to ship later (intertemporal substitution), discussed above.

4 Estimation and Results

We next discuss how we take our framework to the data, as well as the results from our estimation procedure. Section 4.1 covers the port technology parameters; Section 4.2 the demand parameters; and Section 4.3 the cost of infrastructure.

4.1 Port Technology Estimation and Results

Service Time We specify port service time as a general CES specification where,

$$T_{jt} = \omega_{jt} \left(\alpha \left(\frac{L_{jt}}{s_{jt}} \right)^{\eta} + (1 - \alpha) \left(\frac{c_{jt}}{s_{jt}} \right)^{\eta} \right)^{-\frac{1}{\eta}} \tag{8}$$

This encompasses both Cobb-Douglas ($\eta \to 0$) and Leontief ($\eta \to -\infty$) as special cases. Furthermore, following Arellano and Bond (1991), we let port productivity satisfy, $\log \omega_{jt} = \log \omega_j + \theta_\omega t + \nu_{jt}$, where ν_{jt} is drawn iid with zero mean.

Following the usual concern when estimating production functions that inputs are not orthogonal to realized productivity, we introduce instruments, z_{jt}. First, we use local wages in related occupations as an instrument for labor. In particular, wage changes in occupations such as truck drivers, construction equipment operators, and taxi drivers, the most common substitute occupations for longshoremen (as indicated by occupational transition matrices), change their outside options and thus affect port employment. We construct this instrument using wage data from the Occupational Employment and Wage Statistics of the BLS and occupational switching probability data from the Current Population Survey (CPS) of the US Census Bureau. As additional shifters for a port's labor force, we use changes to the labor force and to

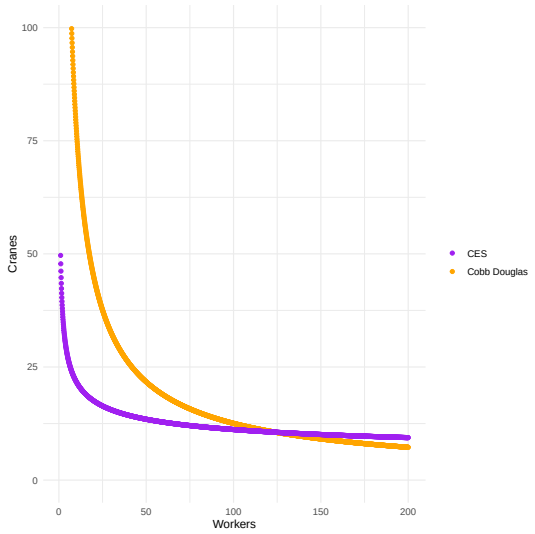


Figure 4: Port service time. The figure plots the isoquants associated with the estimated service time CES specification (8) (for the average service time); it also plots the corresponding isoquant if we were to estimate a Cobb-Douglas specification (i.e. $\eta \to 0$).

the fraction of the population above 65 years old in the port county. Finally, we include lagged inputs as instruments. We leverage these instruments to estimate equation (8) using a GMM estimator based on the moments $\mathbb{E}(\nu_{jt} z_{jt}) = 0$.[28]

As shown in Figure 4, our results suggest that labor and cranes have low substitutability, as the shape of the isoquant is more similar to that of a Leontief (the estimates are in Table 8 in Appendix A).[29] Moreover, we find evidence of substantial productivity dispersion: as shown in Table 7 in Appendix A, there is sizable heterogeneity across ports in the US, even within the same region. The same table shows that our port productivity estimates are consistent with anecdotal evidence that ports that have lower tide range are more productive. Moreover, ports on rivers and ports with man-made breakwater structures are less productive than natural coastal ports.

Port Capacity In our setup, effective port capacity, K_j, is the maximum number of ships that a port can handle at a time. Due to our granular port call data, we can measure K_j directly, instead of using other common but poor proxies, such as the number of berths. For every port we consider the number of ships simultaneously being serviced, conditional on at least one ship waiting outside the port. We

[28] Since we have annual data on labor and cranes, and our instruments are also annual, we estimate equation (8) at the port-year level. In particular, we residualize, separately for each port, the observed service times on ship size and commodity carried and use its yearly average for estimation.

[29] As is typical in most of the literature, we do not allow the shape of the production function to vary over time or across units. Using detailed data on four large US ports (Oakland, Los Angeles, Tacoma and Seattle) from the Pacific Maritime Association, we find supportive evidence for this assumption: the share of crane operators over total labor is approximately 15% and varies little both over time and across these ports. This suggests that there is not much heterogeneity over the technological relationship between workers and cranes across ports. For more details see Figure 17 in Appendix A.

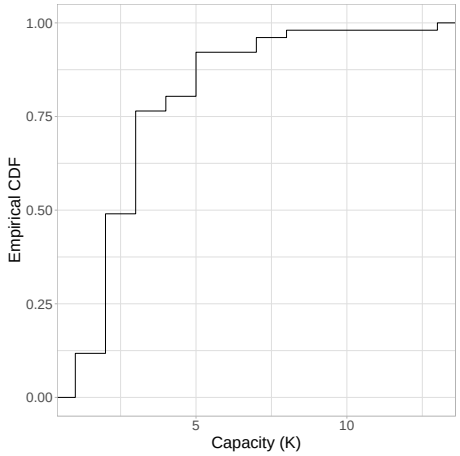


Figure 5: Port capacity. This figure plots the empirical cumulative distribution of ports' effective capacity K_j, measured using AIS data, as the maximum number of ships that can be handled simultaneously, conditional on at least one ship waiting in anchorage.

condition on ships waiting, since this implies there is no space inside the port and therefore we avoid underestimating port capacity. We then take the maximum (90th percentile) of this measure over the sample period and this pins down K_j.

Figure 5 plots the empirical distribution of recovered capacity across ports. Most ports have low capacity: 75% of ports can service at most 3 ships simultaneously. The largest ports, however, have substantially higher capacity; for instance Houston (TX) and Baltimore (MD) can handle up to 8 and 13 ships respectively.

The port's effective capacity K_j is a function of its physical infrastructure, such as the length of its berths and storage space. Estimating the capacity function is crucial when we later gauge the cost of increasing K_j, derived from dredging and acquiring land.

We assume that capacity is given by a Leontief function: a ship requires both space to park (length of berth), as well as the land space to (un)load the cargo, and it is unlikely that the two are flexibly substitutable. Therefore we have,[30]

$$K_j = \min\left\{\frac{B_j}{\kappa_1}, \frac{A_j}{\kappa_2}\right\} \tag{9}$$

where B_j is port j's total length of its berths, A_j is its acreage of storage space, and κ_1, κ_2 are parameters.[31] In particular, κ_1 and κ_2 are technological constants determining, respectively, the mileage of berths and

[30] We omit time variation, as these variables exhibit very little variation in our sample.

[31] Since our measure of length of berths or storage likely includes those for containers, we multiply berth length and storage space with the fraction of bulk ships that the port handles. We collect the data on the percentage of bulk ships served manually from MarineTraffic.

acreage of storage required to handle a ship.

An implicit assumption here is that the parameters κ_1, κ_2 are the same across ports; this is consistent with data: Figure 18 in Appendix A suggests that the length of berths per ship, as well as the storage space per ship are roughly similar across ports. With that in mind, we use our infrastructure data to calibrate these parameters. We set κ_1 equal to the median number of miles of berth per ship, B_j/K_j, and κ_2 equal to the median storage acreage per ship, A_j/K_j.[32]

The Leontief specification proves a very good fit. To showcase this, we also estimate a CES specification (which nests Leontief) for the capacity function. As shown in Figure 19 in Appendix A, which plots the isoquants for $K_j = 1$, the CES ends up being very close to a Leontief. This is consistent with our intuition that length of berths and storage space are not substitutable and are needed in fixed proportions.

4.2 Port Demand

We estimate the demand model (4) from the data on port calls. One complication is that this dataset does not contain information on shipments' US locations d. This is an important impediment to measuring substitution across ports. For instance, consider a shipment that departed from China and unloaded at the port of Long Beach (CA). The shipper will react differently to congestion in Long Beach if the cargo's destination is California, in which case they might prefer to tolerate the congestion or perhaps switch to a nearby port, compared to the case where the destination is in the Midwest, in which case they might decide to opt for a farther away, but well-connected, port. Our data though only reports that the ship unloaded from China to Long Beach, with no information on the shipment's final destination within the US. Similarly in the case of exports, we may observe that a shipment loaded in Houston (TX) and traveled to Europe, but we do not know where in the US it was originated.

In other words, instead of observing $\Pr(j|f,d,t)$, what we observe from port calls is,

$$\Pr(j|f,t) = \sum_d \Pr(j|f,d,t)\, m_{fd} \tag{10}$$

which is obtained by marginalizing $\Pr(j|f,d,t)$ given in (5) by the fraction, m_{fd}, of shipments between

[32] In equilibrium, since A and B are costly to expand, we would expect to have, $K_j = A_j/\kappa_2 = B_j/\kappa_1$ for all j. In other words, we implicitly assume that a port would not invest in storage space or berth length that will not be used. In practice, our measurement of storage space and berth length are noisy since they are hand-collected. Therefore, we observe $\tilde{A}_j = \kappa_2 K_j + \epsilon_j$ and $\tilde{B}_j = \kappa_1 K_j + \eta_j$,where (ϵ_j, η_j) are measurement noise. Our estimator essentially assumes that the errors (ϵ_j, η_j) have median equal to zero. Under this assumption, we recover the parameters via a least absolute deviations (LAD) regression, which is robust to outliers and consistent when errors have zero median. We find that $\kappa_1 = 0.26$, which is 1.5 times the length of the average ship, and $\kappa_2 = 87.1$.

any domestic US location d given the foreign location f.

To tackle this issue, we augment our data with the FAF Regional Database, which contains information on freight movements between US and foreign regions. In particular, the FAF reports annual trade flows between US regions, such as states or metropolitan areas, as well as flows between these regions and broad foreign locations (e.g. flows from California to Southeast Asia or Europe). Putting the two data sources together provides two different aggregations of the unobserved choice probabilities σ_{jfdt}; the first, from port calls, is given in (10) and integrates over domestic locations, d; the second, from the aggregate trade data in FAF, integrates σ_{jfdt} over time, t.

A second complication of our framework is the endogeneity of time at port. Indeed, total time at port may be correlated with other time-varying characteristics of the port captured by ξ_{jft}. For example, an increase in demand driven by higher ξ_{jft}, by construction, leads to higher time at port, TT_{jt}, as more ships now choose port j. For this reason, it is necessary to use an instrument for TT_{jt}. To do so, we construct a measure of unexpected disruptions in port j's service operations at t, which increase the total time a ship spends at port j. These disruptions directly affect TT_{jt} and, through the increased time at port, also reduce demand for port j as fewer shippers now choose this port; at the same time, these disruptions do not directly affect how many ships arrive at port.[33]

To construct the unexpected disruptions in port service operations, we employ the residuals from a regression of service time on a number of controls, including port and time fixed effects, as well as ship size and the number of ships simultaneously serviced; these residuals are meant to capture unexpected disruptions in service, such as crane breakdowns.[34] In Table 9 in Appendix A, we regress time at port on the instrument (akin to a "first stage"); the results indicate that the instrument is strong.

The Estimator The parameters we want to estimate are $\theta \equiv \left\{\beta_F, \beta_D, \beta_T, \beta_P, \gamma_t, \gamma_f, \gamma_{l(j)}\right\}$. Our estimation approach relies on the IO literature (Berry, Levinsohn, and Pakes, 1995); and micro-BLP (Berry et al., 2004) in particular.[35]

[33]One may also worry that prices are also endogenous and thus require an instrument. We are less concerned about prices than time at port. First, as discussed in Section 2.2, prices exhibit no time series variation (see Table 3 in Appendix A). Second, when considering the cross-section, we find that port prices do not seem to be explained by port characteristics (such as port location measured by the distance from foreign locations, service time, capacity, wait time), nor do they seem to respond to characteristics of "competing" ports ("BLP instruments").

[34]In Brancaccio et al. (2025b) we use data from detailed logs from port calls that outline all the activities each ship undertook during a port call. Some of these activities are consistent with unexpected service disruptions, such as the exact timing of the ship's arrival (e.g. whether the ship arrived at the port during the night, the weekend or a holiday), strikes, repairs, gear breakdown, weather. One can think of our instrument as a way to aggregate all such disruptions.

[35]In our context, a market consists of a month-foreign location; we have a total of 560 such pairs and a choice among 51 ports. We treat the shipment's US location d as the unobserved shipment's type. We have eight foreign locations (SW & Central Asia, SE Asia & Oceania, Rest of Americas, Mexico, Europe, Eastern Asia, Canada, Africa) and ten domestic

We first use the FAF data to recover the weights m_{fd}, as well as the market size M_f, which we measure using the maximum observed yearly flow in FAF from/to foreign location f.[36]

A key insight of Berry, Levinsohn, and Pakes (1995), building on Berry (1994), is that, given a value for the parameter β_D and weights m_{fd}, there is a unique vector of δ_{jft} that guarantees that the predicted market shares, $\sum_d \sigma_{jfdt} m_{fd}$, where $\sigma_{jfdt}(\theta)$ is given by (5), are equal to their empirical analogue, i.e. the observed port choices in the port-call data, given by $P(j|f,t)$. In other words, the equations, $\sum_d \sigma_{jfdt} m_{fd} = P(j|f,t)$, can be used to solve uniquely for the δ_{jft}, given the value of β_D, and weights m_{fd}. The recovered δ_{jft} is linear in the remaining parameters of θ, so that given β_D, we can recover the remaining parameters, including the unobserved port characteristics, ξ_{jft}, by regressing δ_{jft} on the covariates in (6).

In order to recover parameter β_D, we minimize a set of moments that are a function of the recovered δ_{jft} and ξ_{jft}, which as discussed above depend on the value of β_D. The first set of moments leverages exclusion restrictions of the form,

$$\mathbb{E}\left[z_{jt}\xi_{jft}\right] = 0$$

where z_{jt} is the instrument for time at port described above, resulting in the moments,

$$m_1(\theta) = \sum_{f,t}\sum_j z_{jt}\xi_{jft}(\beta_D) = \sum_{f,t}\sum_j z_{jt}\left(\delta_{jft}(\beta_D) - \beta_F \text{dist}(f,j) + \beta_T TT_{jt} + \beta_P p_j - \gamma_t - \gamma_f - \gamma_{l(j)}\right)$$

Then, following micro-BLP, we augment the estimator with "micro-moments" based on the FAF Regional Database. In particular, to strengthen the identification of β_D, we construct a second set of moments, $m_2(\theta)$, that match the following object to its theoretical counterpart:

$$\frac{1}{N_{\text{year}}}\sum_{\text{year}=1}^{N_{\text{year}}} P(d|j,f,\text{year})$$

In words, we measure the average fraction of shipments over time whose domestic location is d, given their port j and foreign location, f.[37] For instance, out of the shipments from South East Asia that are

locations (East North Central, East South Central, Mid Atlantic, Mid South Atlantic, Midwest, Mountain North, Mountain South, New England, Pacific, South Atlantic).

[36] It is worth noting that our measure of market size, M_f, is likely underestimated since it relies on the maximum number of shippers observed; it thus ignores potential shippers that have not yet engaged in imports/exports in this market.

[37] Since this quantity can be noisy when the port j is small or foreign location f is a smaller trading partner, we focus on

discharged in Long Beach, this moment measures the fraction whose final destination is California. Thus, our objective function is,

$$arg \min_{\theta} m(\theta)' W m(\theta)$$

where

$$m(\theta) = \begin{bmatrix} m_1(\theta) \\ m_2(\theta) \end{bmatrix}$$

and W is a weighting matrix. We use the state of the art PyBLP package (Conlon and Gortmaker (2020)).[38]

Finally, as described in Section 3.2, we enrich our demand model with a time-series process, given in equation (7), describing the evolution of the aggregate demand shocks γ_t, in order to capture long-term returns from investment.

To do so, we leverage historical macro data. In particular, once we obtain the month fixed effects γ_t we project them on a set of macroeconomic variables y_t, including commodity prices for coal, agricultural products, and metals collected from the IMF:

$$\gamma_t = \rho y_t + \epsilon_t \tag{11}$$

The benefit of introducing the observed macro series, y_t, is that we can now use a much longer time-series (1992-2023) based on these variables to estimate the properties of the macroeconomic conditions, $\hat{\gamma}_t \equiv \hat{\rho} y_t$. Indeed, given the enormous observed volatility, a long-enough sample, that includes several booms and busts, is needed to accurately estimate the demand process.

Once the parameters ρ are estimated, we obtain the parameters in (7) via Maximum Likelihood.

a subset of regions f and ports j. Namely, we match moments for $f \in \{\text{Canada, Eastern Asia, Europe, SE Asia }\}$, and five of the biggest ports in the data: Seattle, Baltimore, Houston, Los Angeles-Long Beach, Hampton Roads. Nonetheless, the results are remarkably robust to the choice of the specific moments to match. Note that in some cases FAF reports a more aggregated version for the point of entry, in which case we have data for a group of ports; for instance LA and Long Beach are lumped together.

[38] Our data features zero market shares for some markets (foreign location and month, ft). To deal with this, we drop ports that have fewer than 10 ships per year to/from f . Results are generally robust to variations of this rule such as lower threshold to drop ports, or different measures of the outside good (e.g. based on FAF vs. port call AXS data).

Demand Estimates	
Average time at port	−0.43
	(0.18)
Port price per ton	−0.33
	(0.006)
Foreign distance (log days of travel)	−2.4
	(0.19)
Domestic distance (log days of travel)	−1.52
	(0.03)
Foreign location FE	Yes
Month FE	Yes
Instrument	Unexpected disruptions to port operations

Table 2: Demand estimates. This table reports the estimated parameters of the payoff function (4).

4.2.1 Results

Table 2 presents the estimated parameters.[39] We begin by computing the "value of time" implied by our estimates, which is measured by the ratio β_T/β_P and captures the amount carriers are willing to pay in order to reduce time at port by one day. We find a value of time of \$1.29 per ton of deadweight, hence $\$64,425$ for the average ship.[40] This estimate for the value of time is substantial, especially compared to the direct cost of $\$14,000$ that the shipper needs to pay to hold the ship for an extra day. The difference between our estimate of value of time and this direct cost reflects the cost of port disruptions as they propagate through the supply chain.

In addition, we find that the elasticity with respect to time at port is large, as shown in Figure 6, suggesting that any disruptions affecting time at port are likely to have a large impact on demand. Moreover it is important to note that there is large heterogeneity in the elasticity across space which we return to below.

Furthermore our estimates suggest that, in addition to time at port, a port's geography and its

[39] In our empirical specification, we proxy for expected time at port, TT_{jt}, with a four-month rolling average of time at port (measured in days). As we discuss below, our empirical findings highlight that the elasticity is substantial, suggesting that shippers indeed react to past congestion. Moreover, we measure distances in days. Days of travel between the foreign location f and port j are computed as maritime distance between port j and the three largest ports in f divided by speed (15 knots per hours). Days of travel between the domestic locations d and port j is computed as miles traveled divided by speed (which is set equal to 65 miles per hour for 12 hours per day).

[40] As a robustness, we have re-estimated demand restricting the sample to the smaller ship sizes (Handysize and Handymax) for which we have enough observations. We find that the value of time is in the same ballpark. The commodity shipped is another dimension of heterogeneity. However, once we control for the origin, destination, and whether the ship is loading or unloading, we are left with virtually one commodity handled. This suggests that the heterogeneity across commodities is subsumed by the heterogeneity we already control for and, therefore, controlling for commodity in the demand specification does not seem first order.

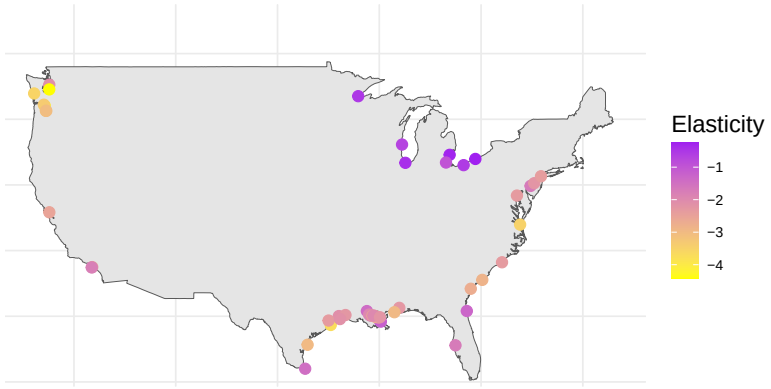


Figure 6: Demand elasticity. This figure plots the average elasticity with respect to time at port for each US port.

characteristics also affect demand. In particular, our demand model allows for different forces that shape substitution patterns across ports, such as the distance to the foreign and domestic locations, as well as port characteristics reflecting factors such as better services offered or better connectivity.

It is useful to understand the importance of geography vs. time at port. Figure 7 plots each port's geographic attractiveness δ^{geo} , measured by the term $\beta_F \text{dist}(f,j) + \beta_D \text{dist}(j,d) + \gamma_{l(j)} + \xi_{jft}$, averaged across domestic and foreign locations, as well as time t, and normalized to be between zero and one. Note that δ^{geo} depends purely on ports' geographic attributes, and purges variation coming from the ports' different congestion levels. The figure reveals that the most geographically attractive ports are ports in the Gulf, such as Houston (TX), that serve the flow of raw materials from the Mid Atlantic range and Southern states to Europe, Africa, and Latin America. Similarly, ports in the East Coast, such as Hampton Roads (VA) or Baltimore (MD), and the Northwest Coast, such as Portland (OR), are also highly desirable, as they trade in raw materials with Europe and China respectively.

Do traders use the geographically most convenient port? Interestingly, only 48% of market-shipper pairs choose the port that is the most desirable geographically, as defined by δ^{geo} of that particular (f,d) pair. This suggests that time at port acts as a barrier to trade, pushing traders to sacrifice geographic convenience in order to avoid port congestion.

We further explore this by examining how shippers respond to an increase in time at port. Figure 8 plots the fraction of shipments that are diverted to a port's 5 closest neighboring ports, conditional on shippers choosing to switch. On average 34% of diverted trade is diverted locally. Again, however, there is large heterogeneity across ports. In general, the figure shows that for ports in the Great Lakes substitution tends to be local, while the reverse is true for ports in the Gulf.

It is instructive to compare two ports displaying these opposite patterns, Corpus Christi (TX), where

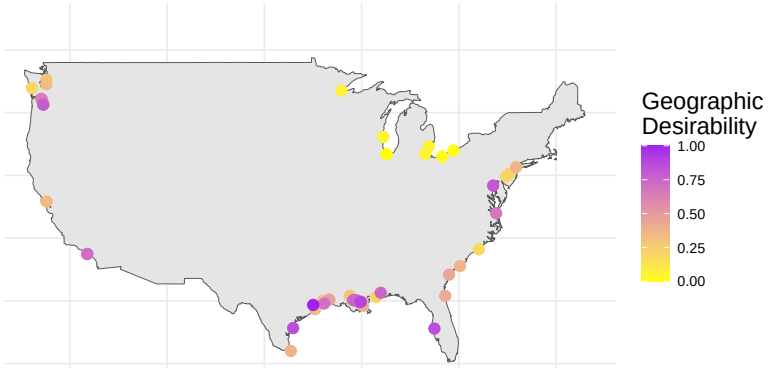


Figure 7: Geographic attractiveness. This figure plots the variable δ^{geo}, measured as $\beta_F \text{dist}(f,j) + \beta_D \text{dist}(j,d) + \gamma_{l(j)} + \xi_{jft}$, for each US port, averaged across domestic and foreign locations, as well as time t. Domestic and foreign locations are weighted by the market size of each (f,d) pair. This captures the geographic attractiveness of each port.

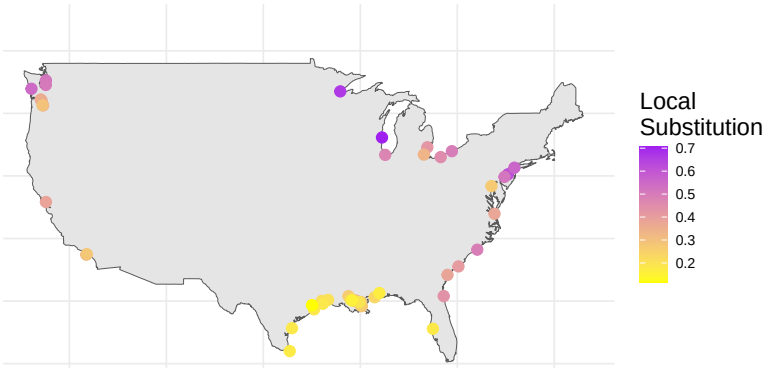


Figure 8: Local substitution. For each port j this figure plots the fraction of shipments through j that is diverted away from j to a local port (i.e. to one of the five closest ports), conditional on shippers choosing to switch following a small increase in time at port.

18% of trade is diverted locally, and Cleveland (OH), where 45% of trade is diverted locally. Corpus Christi (TX) is very attractive for a number of (f,d) pairs that are often far away, mostly serving trade between Europe and Latin America, and the Atlantic Coast and Midwest. Therefore, when Corpus Christi (TX) is congested, these importer-exporter pairs might respond by substituting to Hampton Roads (VA), a large port on the East Coast, well connected with these origins and destinations. In contrast, Cleveland (OH) mostly serves local shippers, i.e. it is attractive mostly to local (f,d) pairs, operating between Canada and the Midwest. Naturally, these trades have less flexibility and tend to rely on other local ports.

These substitution patterns are helpful in understanding the large geographic heterogeneity in elasticities with respect to time at port, shown in Figure 6. For instance, shippers in the Great Lakes are fairly far away from the most desirable ports and when they substitute they tend to do so locally. Having said that, precisely because there are no good options nearby, the estimated elasticities of the Great Lakes

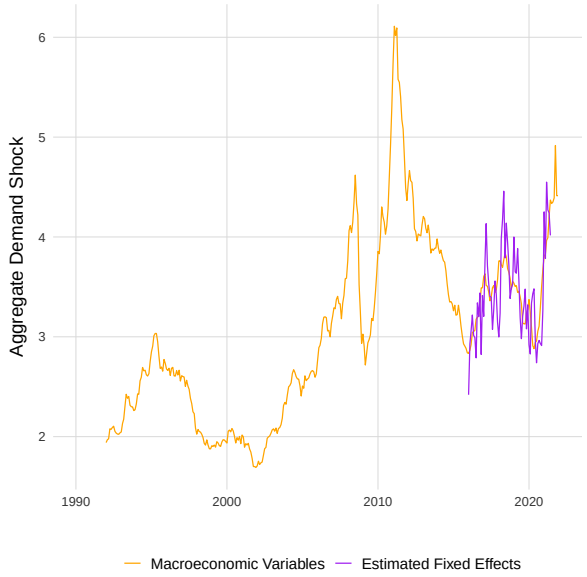


Figure 9: Macroeconomic volatility. This figure plots the estimated fixed effects, γ_t, as well as the time fixed effects predicted by macroeconomic variables, $\hat{\gamma}_t \equiv \hat{\rho} y_t$. The macroeconomic variables y_t include prices for coal, agricultural products and metals. See Table 10 in Appendix A for details.

are the lowest in the country. Conversely, ports in the Gulf or the East Coast have several substitutes in a wider geographical range, so they do not substitute locally and they feature high elasticities. Finally, ports in the Northwest Coast are close to some of the most attractive ports in the country, so they have both high elasticities and a much higher propensity to substitute locally when they do substitute.

Finally, Figure 9 plots the estimated time fixed effects, γ_t, which capture macroeconomic volatility, while Table 10 in Appendix A reports the estimated coefficients for the AR(1) process for $\hat{\gamma}_t \equiv \hat{\rho} y_t$ given in (7), as well as the estimated parameters ρ in (11) that allow us to extrapolate to a longer time-series.[41] Besides demonstrating that the fit of our time series process is quite good, Figure 9 also reveals the severe volatility of our setting. A key feature of transportation markets is that demand (captured here by γ_t) exhibits large booms and busts, driven by macroeconomic fluctuations. The COVID-19 disruption is clearly visible and large. Interestingly, however, it is not incomparable to other recent disruptions. For instance, it is comparable to the dramatic boom and bust of 2006-2008, as discussed in Section 2.1. The combination of this macroeconomic volatility with the capacity constraints of the transportation sector, renders transportation markets exceptionally volatile (Brancaccio et al., 2025a, 2025b). It is therefore necessary to account for this volatility when computing the welfare gains for a long term investment in capacity expansion like the one studied in this paper.

[41] Note that this exercise is predictive, and there is no economic interpretation of these coefficients, which capture both demand and supply effects. For instance, when food prices go up, supply for exports increases, while demand for imports falls. Hence, there is no obvious expected sign for the food prices.

4.3 The Cost of Expanding Port Capacity

As discussed in Section 2.2, in order to recover the cost of port infrastructure we rely on cost estimates from the US Army Corps of Engineers (USACE), as well as land costs. Authorities rely on the same information from the USACE to make investment decisions.

Consider the cost of installing an additional unit of capacity, i.e. the cost of increasing port j's capacity from K_j to $K_j + 1$, so that the port can handle one additional vessel. The cost for port j is equal to,

$$C_j^d + C_j^l + C_j^b$$

where C_j^d is the cost of dredging a section of ocean or riverbed at port j sufficient to accommodate the ship; C_j^l is the cost of purchasing a parcel of land bordering port j to serve as a storage facility for goods transiting the new berth; and C_j^b is the cost of constructing a bulkhead to support the new berth.[42] We discuss each term in turn. Appendix C provides the details.

Accommodating an additional vessel requires dredging a portion of the seafloor with the volume determined by the vessel's size and the port's current depth; the associated cost, C_j^d, crucially depends on what share of this volume requires blasting through rock. For each berth in our dataset, the USGS provided us meter-by-meter estimates of underwater wave velocities, computed based on Boyd (2020).[43] From these, we calculate the depth at which the seabed switches from rock material to sediment, whereafter drilling and blasting must be used. In other words, we obtain an estimated depth to non-dredgeable rock material for each berth.

We estimate the cost function of dredging rock and non-rock material from cost estimates obtained by the USACE. In particular, we are able to use data from two expansion projects: Delaware in 2010 and Boston in 2017. For these projects, the USACE provides detailed estimates of predicted dredging costs and material quantity for 78 berth-depth combinations. We estimate a linear function of costs on total quantity dredged (for both rock and non-rock; the total cost is the sum of the two).

Next we turn to the cost of storage, C_j^l. This is equal to the price of industrial land around port j, p_j^A, times the acreage A_j required to accommodate the additional ship, i.e., $C_j^l = p_j^A A_j$. To obtain land prices around ports, we first use OpenStreetMap to isolate industrial and commercial land in proximity to each berth. Then, we find the corresponding value of this land from data made public by Nolte (2020). We obtain the acreage required to accommodate the additional ship from the Leontief capacity function

[42] A bulkhead is a vertical shoreline stabilization structure that primarily retains soil, and provides protection from waves.
[43] We are most thankful to Oliver Boyd for sharing this data with us.

(9).

We also need a measure for the cost of bulkhead construction, C_j^b. The USACE reports that the cost of erecting a new bulkhead at Charleston (SC) amounted to $22 million, or $6,600 per foot of berth. We use this number as our baseline and perform robustness around it.

Finally, we assume that, when investing in port infrastructure, the port increases labor and cranes in order to maintain the service time constant and we add the cost of purchasing cranes and hiring workers to the total expansion cost.[44] To recover the number of cranes and workers required, we use the CES service time specification (8). We calibrate the cost of a crane to 20 million USD, based on estimates from the port of Savannah (GA), and the annual wage at $\$60,000$ (US Department of Labor).

Overall, total costs are 115 million USD on average with a fair amount of geographical dispersion (standard deviation equals 98 million USD). Land costs account for about 49% of total costs on average, and dredging costs for 26%. Land costs account for an even larger share of total costs in urban areas such as Portland (OR) (87%), Seattle (WA) (85%), and New Orleans (LA) (80%).

Finally, let us note that there are two caveats to this data. First, they rely on two port projects (albeit with several berths each), as these are the only projects reporting detailed information on costs and quantities excavated.[45] Second, they rely on engineering estimates, while economists often worry about other types of costs that are relevant for investment decisions (such as opportunity costs, disruptions to port operations, political constraints etc.). Importantly, however, policy makers rely on the same cost data, which they benchmark against a (local) benefits analysis of increased port throughput; one can think of our exercise as comparing the same costs to a model for "benefits" that takes into account dynamics, demand feedbacks, and spillovers across ports.[46]

[44] This is done in order to recognize that adding a unit of capacity will not be valuable if the port does not have the resources to man it. This requires both cranes and personnel. So our presumption is that acquiring a new unit of capacity must come hand in hand with these additional costs. We ran a robustness check where we exclude the cost of cranes and workers from the cost of investment. We find that the these additional costs account for 9% of total costs, with the typical port requiring four additional workers. Removing these costs changes the investment decision for only two ports—both with marginal net gains in the baseline.

[45] Reassuringly, the estimated cost function across the two projects is very similar; if we only use each project in isolation to estimate the dredging cost function, we obtain very similar results.

[46] An alternative approach would be to estimate the cost of infrastructure from port investment decisions, following the IO literature. There are two important issues with this approach that made us uncomfortable using it. First, this requires a model of port behavior; yet it is quite difficult to gauge what the objectives of port authorities are: as an example, as we show in Table 3, port prices are fairly flat over time and do not seem to change as demand fluctuates over time. Second, investments in port expansion are infrequent in the US.

5 Welfare Analysis of Infrastructure Investment

Billions are spent every year from both public and private actors for investment in port infrastructure. What are the welfare gains of these investments? When and where is port investment justified? How should a budget be allocated? How frequent do macroeconomic shocks need to be to warrant investment? We now use our estimates for port technology, demand for port services, and costs of infrastructure to answer these questions and provide a comprehensive welfare analysis. We first calculate the gross and net returns from increasing capacity at each port by one unit, akin to a cost-benefit analysis (Section 5.1); this also allows to better understand the mechanisms through which infrastructure investment affects welfare and trade. We then explore the role of macroeconomic volatility (Section 5.2). Finally, we consider a planner problem of simultaneous investment in multiple ports (Section 5.3).

5.1 Gross and net returns of investment

We begin by computing the aggregate welfare gains to an increase in the capacity of the US port system by one slot. In other words, a single port can now handle one additional ship; this is similar to the approach followed in Allen and Arkolakis (2022). To put this in context, this increase corresponds to approximately a 1% increase of the total US capacity. In practice, we increase K_j to $K_j + 1$ for a given port j and then we repeat this exercise for all ports, one port at time. We use our model to simulate a long-run counterfactual time series of 30 years. As discussed above, we assume that along with K_j, port j's labor and cranes also increase so as to keep its (average) service time constant. Note that since many small ports can handle only 1 or 2 ships at a time this increase in capacity can be substantial. We describe how these simulations are conducted in Appendix D.

Below, we first outline the economic mechanisms that are set in motion when the capacity of a port increases. We then calculate the welfare (and trade) gains, and thus obtain the "benefits" from these investments. Finally, we compare these welfare gains to the infrastructure costs to obtain a measure of net returns and complete the cost-benefit analysis.

Mechanisms We first describe the impact on the "treated" port, and then explore what happens to all other ports.

When capacity K_j increases at a **treated port** j, its wait time declines, all else equal. In turn, the reduction in congestion makes the port more appealing and attracts ships both from competing ports and from the outside good. It is important to note that the latter creates a net increase in total trade, unlike

the ships' reallocation, which is a transfer of trade from other ports to port j.

The endogenous response of demand at port j dampens the original fall in congestion. Indeed, although congestion initially declines because of the higher capacity, the resulting higher demand mitigates this decline, boosting congestion and time spent waiting. In other words, while increasing K_j to K_j+1 reduces port j's wait time, this reduction is mitigated by the increased arrival of ships that respond to this change. This is reminiscent of the so called "Iron Law of Congestion" (Downs, 1962, 1992; Duranton and Turner, 2011) often evoked in road construction: as road capacity increases, new drivers use the road, mitigating any relief in congestion. In our context, this inflow of new customers is socially desirable as it potentially constitutes new (international) trade that was not taking place before because of port congestion.

In summary, we expect to see a potentially considerable increase in trade at the treated port (depending on the port's attributes, an issue we return to below), but a more modest decline in congestion. Indeed, we find that on average (median) trade increases by 16.30% (11.57%) at the treated port, while time at port declines on average (median) by 7.85% (6.89%).

Now let us turn to what happens to **all other ports**, when capacity increases at port j. In this interconnected system, changing capacity in one port does not occur in a vacuum. Instead, it is bound to have an impact on the entire port system, as suggested by the above discussion. Improving port j does not affect just j but also other ports, whose demand is reduced as port j now offers better wait times. In other words, there is substitution to the treated port. This *substitution effect* implies that the increase in trade at the treated port does not necessarily increase aggregate trade.

To illustrate this feedback effect with an example, we consider an increase in the capacity of Hampton Roads (VA), holding time at port fixed at other ports for now. As shown in the left panel of Figure 10, this change can substantially reduce ship arrivals at competing ports; indeed several ports in the Gulf and East Coast see their demand decline up to 6%, all else equal.

But the story does not end here. As fewer ships choose these other ports, they become de-congested, which in turn endogenously boosts their demand. This feedback effect in neighboring ports creates positive spillovers: investing in infrastructure in port j can decongest not just j, but also a wider set of ports, especially those that are close substitutes of j. We illustrate this in the right panel of Figure 10, which plots the full effect of increasing capacity at Hampton Roads (VA). Interestingly, this feedback effect can be quite large: for several ports, such as Baltimore (MD), the full effect is small, despite a considerable substitution effect. In other words, although a large share of shippers substituted away from Baltimore (MD) towards Hampton Roads (VA), time at port at Baltimore (MD) fell sufficiently to

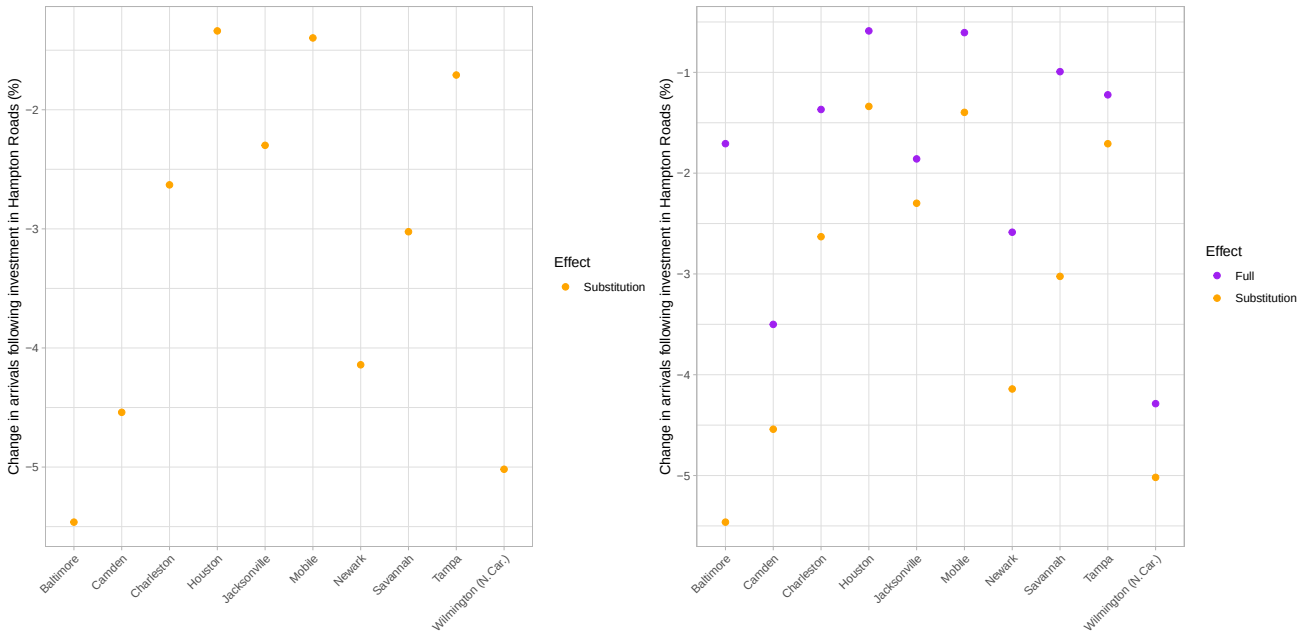


Figure 10: Substitution and full effect of increasing port capacity at Hampton Roads (VA) from K_j to $K_j + 1$, so that it can now handle one additional ship. The left panel plots the change in arrivals at ports other than Hampton Roads (VA), holding time at port at its original level there. The figure depicts the port's top 10 substitutes. The right panel plots the full effect on other ports' demand in response to an increase in infrastructure in Hampton Roads (VA) (i.e. not holding time at port constant at other ports).

re-attract demand to the point that overall Baltimore (MD) did not lose much traffic. In this case, the overall effect on Baltimore's trade is small, as the two forces roughly offset each other.

In summary, at all other ports, we expect to see some decline in trade, as well as some decline in congestion. Indeed, we find that at all other ports on average trade decreases by 0.23%, while congestion declines on average by 0.11%.

The previous discussion has focused on changes in trade flows to different ports. More generally, changes in aggregate welfare reflect increases in aggregate trade, as well as gains stemming from shipper reallocation. For instance, shippers who now choose to use the port of Hampton Roads (VA) instead of Baltimore (MD) might benefit from closer proximity to their final destination, or the lower congestion. We next discuss how investment in infrastructure affects aggregate welfare.

Welfare (and Trade) Gains The key outcome a policymaker is interested is not congestion per se, but rather the gains in trade and welfare that are generated by new infrastructure. As described above, in order to evaluate the welfare impact of capacity expansion in a single port, we need to also take into account the overall effects across all ports. Figure 11 depicts this total change in trade for *the entire* US port system when we increase the capacity of one port at a time; Figure 12 depicts the corresponding

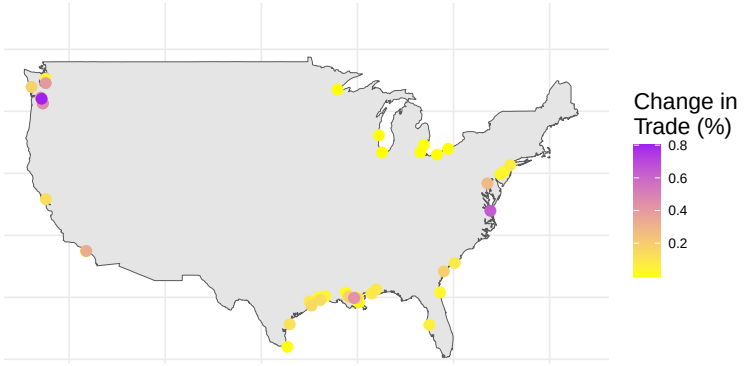


Figure 11: Trade impact of investment in infrastructure. In this figure, we increase port capacity of port j from K_j to $K_j + 1$, for all j, one at a time, so that port j can now handle one additional ship. The figure depicts the total change in trade from increasing capacity in each port j, one at a time.

gross welfare gains to infrastructure investment, which in our setup is the change in *total* welfare from changing a port's capacity. Table 11 in Appendix A contains the change in wait time, trade and welfare from expanding capacity K_j to $K_j + 1$ for every port in our data.

We find that, on average, investing in infrastructure increases aggregate trade by 0.23% and aggregate welfare by 0.14%. Importantly, however, the returns are quite heterogeneous across ports. Indeed, welfare can increase by more than 0.9% when investing in ports such as Hampton Roads (VA) or Destrehan (LA); but they are negligible when investing in other ports, such as the ones in the Great Lakes. This suggests that targeting investment to the "right" port, as opposed to indiscriminate investment, can increase the returns to investment manyfold.

In addition, our estimates suggest that spillovers are large. In particular, if we only allow for the substitution effect, the aggregate increase in trade is 6% lower on average, and up to 23% lower for the ports with the highest gains.[47]

What determines whether a port generates high or low welfare gains? First, size alone does not directly predict welfare gains from infrastructure investment. As an example, Hampton Roads (VA) is the sixth largest port in the US, but is among the ports that deliver the highest gains. To explore the drivers of welfare gains from investing in infrastructure, we consider how these gains depend on both the characteristics of the treated port, as well as those of the set of ports it decongests. Consider again Hampton Roads (VA). Following our discussion in Section 3.1.1, an increase in K_j should have the largest impact on ports that are already congested: investing in infrastructure in ports with spare capacity has little effect; on the contrary, an additional slot in ports that have low capacity relative to their demand, can have a substantial impact, as higher K_j helps absorb the high demand. As shown in the left panel of

[47] As above, to isolate the substitution effect we compute the change in trade holding time at port at other ports fixed.

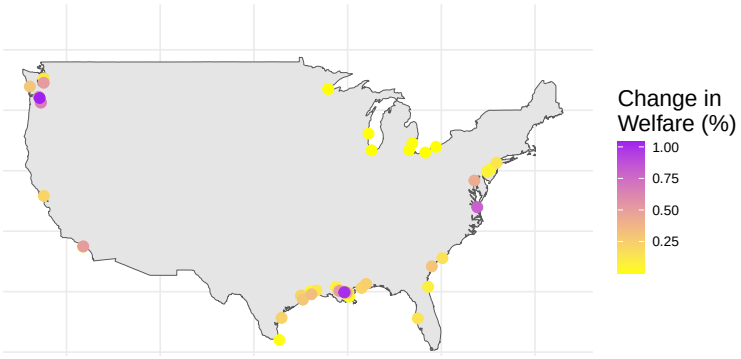


Figure 12: Welfare impact of investment in infrastructure. In this figure, we increase port capacity of port j from K_j to $K_j + 1$, for all j, one at a time, so that port j can now handle one additional ship. The figure plots the total change in welfare from increasing capacity in each port j, one at a time.

Figure 13, Hampton Roads (VA) is indeed a congested port.

In addition, the geographic attractiveness of the treated port is bound to be decisive for its welfare gains. Indeed, investing in ports that are in a convenient location can both attract more shippers from the outside option, and benefit existing shippers that reallocate there. As shown in the right panel of Figure 13, Hampton Roads (VA) is an attractive port in terms of the variable δ^{geo} , which primarily captures the proximity of the port to diverse origins and destinations.

However, looking at the treated port alone does not suffice to break down the overall welfare gains it generates– indeed Hampton Roads (VA) is not the most congested port, nor the most attractive. We must also consider the set of ports that obtain the highest spillovers, i.e. the set of ports that the treated port decongests. In line with our discussion above, if these are ports themselves capable of generating large gains, the overall impact will be higher. Figure 13, demonstrates that the ten closest substitutes of Hampton Roads (VA), are themselves both fairly congested, and geographically attractive.[48]

This intuition holds more generally. As shown in Table 12 in Appendix A, welfare gains are higher for ports that have higher average levels of congestion, captured by the average time that ships have to wait in anchorage, and for ports that have higher geographic attractiveness, δ^{geo} . Moreover, gains increase in the average congestion and δ^{geo} of the treated port's closest substitutes. As discussed in Section 4.2, congestion acts as a barrier to trade; investing in ports that are geographically convenient ensures they are accessible to a larger number of shippers, lowering the transportation costs that they need to shoulder.

The presence of the spillovers that we document raises important policy questions regarding the coordination of investment activities. An implication of our analysis is that unless decisions are made centrally, they are unlikely to internalize the spillovers across ports. While some port investment decisions

[48] We identify these ports as those with the highest cross derivative with respect to time at port at Hampton Roads (VA).

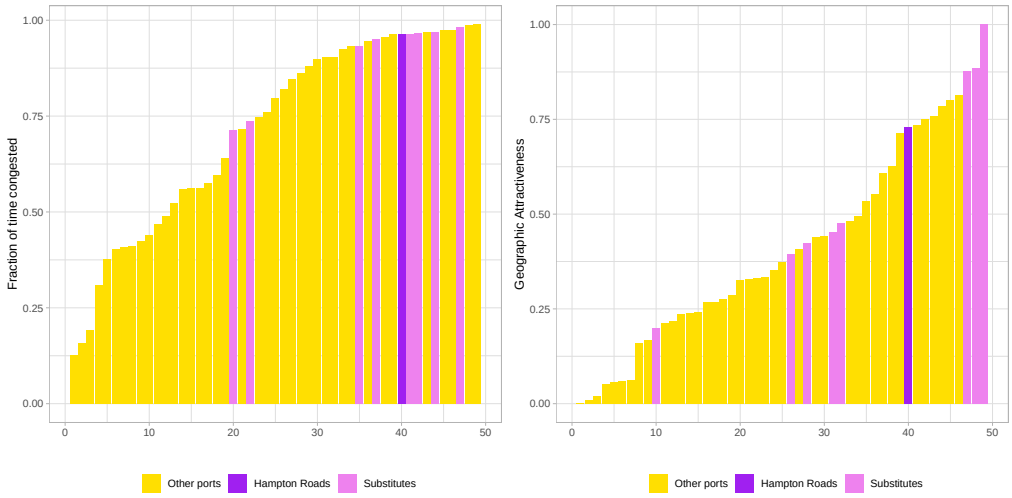


Figure 13: What determines whether a port generates high or low welfare gains? The left graph ranks ports by their queue. Focusing on Hampton Roads (VA), a port with one of the highest returns, we see that it is fairly congested. In addition, its ten closest substitutes are themselves also fairly congested. The right graph ranks ports by their geographic attractiveness, δ^{geo} . Hampton Roads is the most geographically attractive port and its ten closest substitutes are themselves highly attractive ports.

are made at the federal level, most of the funding comes at the private or local (e.g. city, state) level. For instance, the IIJA enacted in 2021 allocates $17 billion for port infrastructure;[49] however this is only a small fraction of the $163 billion to be spent between 2021 and 2025. Given that local authorities seem to care predominantly about the port's local effects, one may be concerned that these spillovers are not fully internalized. Indeed, this lack of coordination and competing incentives might explain why, for some ports the net returns from investment can be sizable.

Net Returns In Figure 14 we compare the welfare gains from increasing K_j to $K_j + 1$, calculated in Section 5, to the corresponding costs, calculated in Section 4.3, and highlight in purple (yellow) the ports for which the former are higher (lower) than the latter. Net returns are positive for 22 out of 51 ports, and are highest for the ports in the Gulf, such as Destrehan (LA), and Convent (LA), as well as for Hampton Roads (VA). Both welfare and costs shape up the net returns: as shown in Figure 20 in Appendix A, although the ports with the highest welfare gains also generate positive net returns, there are several ports, which may be yielding high welfare gains, but also exhibit prohibitive costs, turning net returns there negative. For instance, the ports of Portland (OR) and Long Beach (CA) have some of the highest welfare gains from investment, but are both located in urban areas which are associated with a steep cost of land expansion. The last column in Table 11 in Appendix A contains the rate of return from expanding capacity K_j to $K_j + 1$ for every port in our data.

[49] https://www.maritime.dot.gov/about-us/bipartisan-infrastructure-law-maritime-administration.

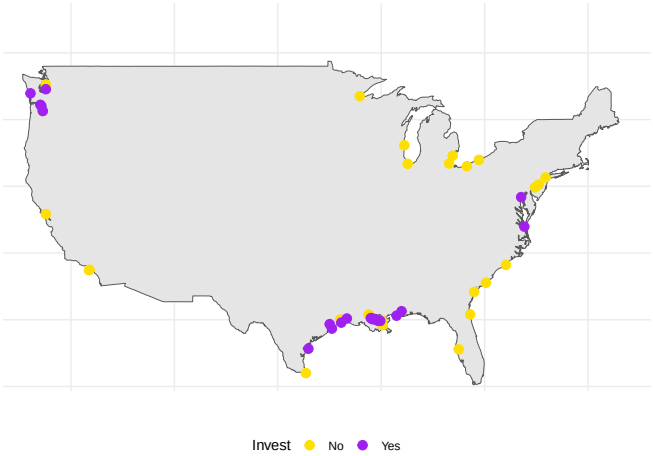


Figure 14: Net returns of investment in infrastructure. In this figure, we increase port capacity of port j from K_j to $K_j + 1$, for all ports, one at a time, so that port j can now handle one additional ship. The graph highlights in purple the ports where the returns from investment are higher than the costs, and yellow for the reverse.

5.2 Macroeconomic Volatility

An important innovation of our framework with respect to the literature is that it explicitly considers congestion dynamics, which illustrates that infrastructure investment is much more valuable when demand surges. We can thus address a timely policy question: how frequent do macroeconomic shocks need to be to warrant investment? In recent years, there is increasing evidence that shocks are becoming more frequent and that uncertainty is rising (see e.g. Bloom, 2009, as well as the literature on the recent supply chain disruptions); at the same time policymakers appear increasingly concerned about resilience.[50]

In Figure 15 we explore net returns from investment in infrastructure as the frequency of macroeconomic shocks gradually rises.[51] Following our discussion in Section 3.1.1, for the average port, welfare gains increase steeply with volatility, and they increase by 75% as volatility doubles, as shown in the left panel of Figure 15. Even more strikingly, at the 90^{th} percentile, welfare gains more than double when volatility doubles. At the other extreme, at the 10^{th} percentile, welfare gains are only modestly affected by the volatility increase.

The right panel of Figure 15 plots the increase in aggregate volatility that makes net returns positive for different ports. As already discussed above, for 22 ports net returns are already positive at the observed volatility. At the other extreme, for a sizable share of ports (35%) net returns do not turn positive even if volatility more than triples; for these ports, investment is unlikely to be warranted even if the policymaker

[50] Indeed, the World Uncertainty Index has roughly doubled in the last twenty years.

[51] In practice, we repeat the long-run simulations under a mean-preserving spread of the aggregate demand process (i.e. we increase the standard deviation, σ_ϵ, of the aggregate demand γ_t). Of course, our estimated net returns would be amplified if the mean of the demand process also were to increase.

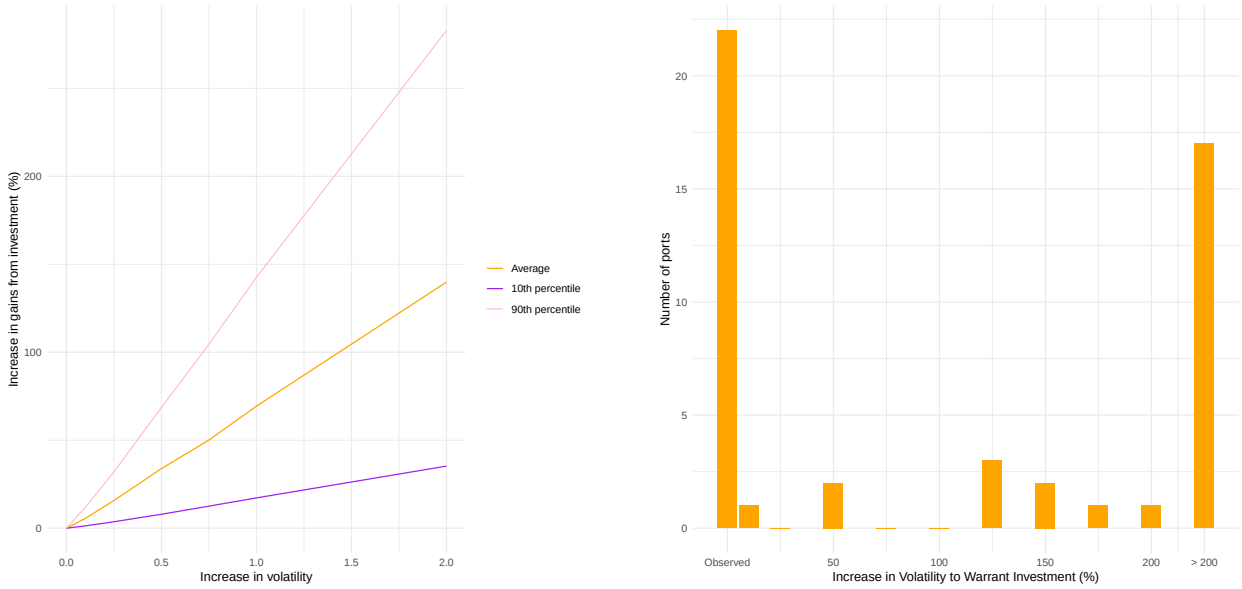


Figure 15: Rising macroeconomic volatility. The left panel plots the welfare gains from investment in the average, as well as the 10^{th} and 90^{th} percentile of US ports, for different percent increases of aggregate volatility. In particular, we repeat the long-run simulations under a mean-preserving spread of the aggregate demand process (i.e. we increase the standard deviation, σ_ϵ, of the aggregate demand γ_t). The right panel plots the histogram of the increase in aggregate volatility that turns net returns positive for different ports.

only cares about resilience. Finally, there is an intermediate range of ports for which some increase in volatility makes investment valuable; for instance, 9 additional ports yield positive net returns if volatility were to triple.

5.3 Investment in Multiple Ports: A Planner Problem

The analysis so far considers welfare gains for each port, when investing at one port at a time. What about simultaneously investing in multiple ports? This planner problem is quite complex. Here, we consider a planner that invests in ports sequentially, as long as net returns remain positive.[52] As shown in Figure 16 which plots the cumulative net returns of investment for each additional unit of capacity, the planner concentrates a large share of their investment in the three ports with the highest individual returns- Kalama (WA), Destrehan (LA) and Hampton Roads (VA)- which receive increases of thirteen, eight and six units of capacity, respectively. In fact, investment in these three ports accounts for 79% of the total budget and generates 96% of the total welfare gains. Overall, this investment program would generate $8.4 billion in total net gains in the form of increased trade through U.S. ports, with a total

[52]In particular, the planner ranks the net returns of all ports. They first add one unit of capacity to the port with the highest returns. They reevaluate the returns, and decide where to invest the next unit of capacity. The planner continues until net returns turn negative. There is, of course, no guarantee that this approximate solution results in a globally optimal allocation of investments. We focus on the ten ports with the highest individual net returns for computational tractability.

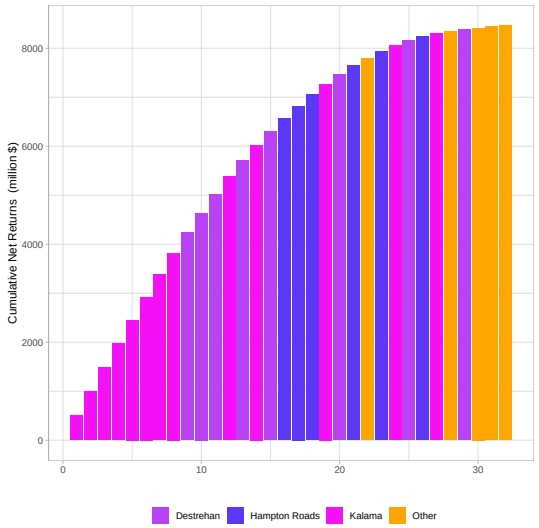


Figure 16: Cumulative net returns from sequential investment in port capacity. This figure considers a planner that invests additional capacity units sequentially until returns turn negative.

budget of $1.6 billion.

It is striking that these ports are each located in one of the three main trading regions of the US: the West Coast, the Gulf, and the East Coast. In other words, rather than investing in a broad set of ports, the planner prefers to concentrate investment to three regional megaports. This reflects substantial economies of scale arising from queueing (see Section 3.1.1): in particular, even if two ports have similar demand and capacity, the randomness in arrivals implies that with some probability one port is backed up, whereas the other has spare capacity. Pooling capacity together allows the planner to avoid these wasteful situations, thereby reducing wait times. Note that a single regional megaport may, however, imply increased distance from the final destination for several (f, d) pairs. The gains in terms of reduced wait time resulting from the economies of scale are substantial enough in this application to outweigh the increased distance. At the national level however, the tradeoff changes and the planner chooses to invest in three regional megaports, rather than a single one. Finally, the above results illustrate once more the costs of decentralized planning in this setting.

6 Conclusion

In recent years, infrastructure investment has been in the pubic spotlight. Ports in particular have undergone a series of major disruptions, from COVID-19, to chokepoint closures. How costly are these disruptions? In this volatile environment, what are the returns to investment in infrastructure? How should a budget be allocated? How frequent do macroeconomic shocks need to be to warrant investment?

Answering these questions requires a clear understanding of congestion formation, shippers' response to shocks, and the interdependencies inherent in the port network. This paper develops simple tools to address these issues. We showcase that disruptions are significantly costly. Yet, for investment to be justified, it must be coordinated and targeted. This is because of the substantial heterogeneities across ports, as well as the sizable spillovers generated by investment. In addition, as these are long-run investments, their returns are crucially affected by macroeconomic volatility. The increased frequency of shocks, or a stronger policy interest in resilience, increase the payoff from investment. Finally, the economies of scale arising from queuing, encourage investment in only a few, select megaports.

Appendix

A Additional Tables and Figures

Dependent Variable:	Dues per Port Call (log)			
	(1)	(2)	(3)	(4)
Deadweight (log)	0.6169***	0.6176***	0.6143***	0.6218***
	(0.0445)	(0.0435)	(0.0437)	(0.0255)
port FE	Yes	Yes	Yes	
loading/unloading FE	Yes	Yes	Yes	Yes
cargo FE	Yes	Yes	Yes	Yes
year FE		Yes		
month FE			Yes	
port-year FE				Yes
N	24,119	24,119	24,119	24,119
R^2	0.687	0.688	0.690	0.693

Table 3: Variation in port charges. This table presents a regression of port prices on ship characteristics, as well as time and port fixed effects.

	log (num)	log (hrs)
log (members)	1.031***	0.980***
	(0.097)	(0.104)
N	81	81
R^2	0.589	0.530

Table 4: Correlation between union membership and annual man-hours for ports on the West Coast using shift-level data from the PMA (Pacific Maritime Association). We regress the number of total workers (log(num)) in shifts in the detailed data from the PMA, as well as total hours (log(hrs)) on the total number of union members. The estimates suggest that not only is the correlation high, but also that the elasticity is equal to one.

	Service time (log)	
# of other ships serviced (log)	0.72^{***}	0.73^{***}
	(0.076)	(0.072)
queue (log)		-0.025
		(0.020)
N	59,576	59,576
Port FE	Yes	Yes
Commodity FE	Yes	Yes
Month-year FE	Yes	Yes
Ship size	Yes	Yes

Table 5: Service time and number of ships at port. In this table we regress port service time on the number of ships simultaneously handled, as well as the number of ships in anchorage, controlling for port, commodity, and month-year fixed effects. Service time per ship depends on how many other ships are serviced at the same time, but not on the queue of ships waiting.

	Total workers per shift
n_{bulk}	31.21^{***}
	(2.49)
$n_{container}$	83.67^{***}
	(1.88)
$n_{container} \times n_{bulk}$	0.14
	(0.42)
N	27,251
R^2	0.86
Port FE	Yes
Year-Month FE	Yes

Table 6: Interaction between bulk ships and containerships. We use shift level data for West Coast ports from the PMA (Pacific Maritime Association) to test for interactions between bulk and containership operations within a port. We find that the total number of workers per shift increases as a higher number of containerships ($n_{container}$) or bulk ships (n_{bulk}) are serviced. However, the interaction is not economically and statistically significant.

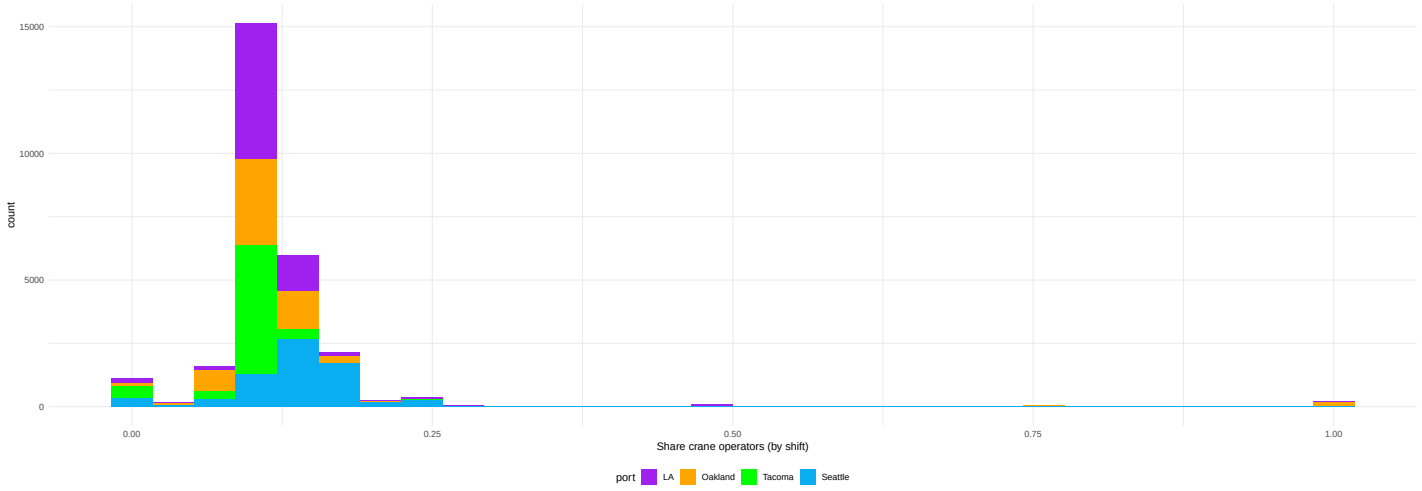


Figure 17: Is crane technology substantially different across ports? This figure uses granular labor data for the major West Coast ports from the PMA (Pacific Maritime Association) and plots the histogram of the share of crane operators with respect to all port workers and illustrates that it does not vary much across ports and over time.

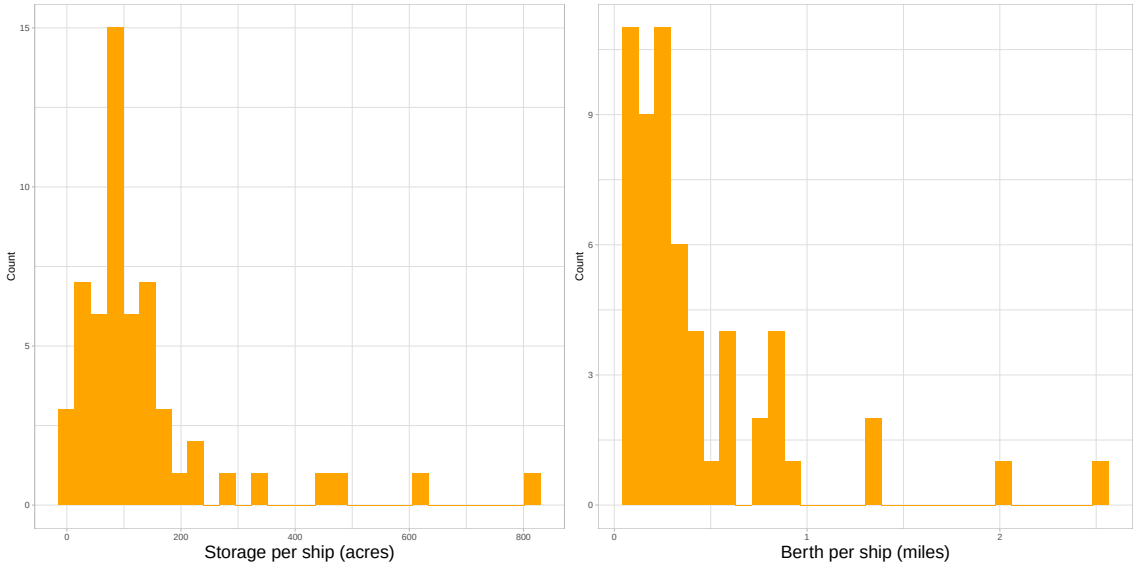


Figure 18: Storage space and berth length per ship. This figure plots the histogram of the ratio of storage space per ship across ports, i.e. A_j/K_j, as well as the ratio of length of berth per ship across ports, i.e. B_j/K_j.

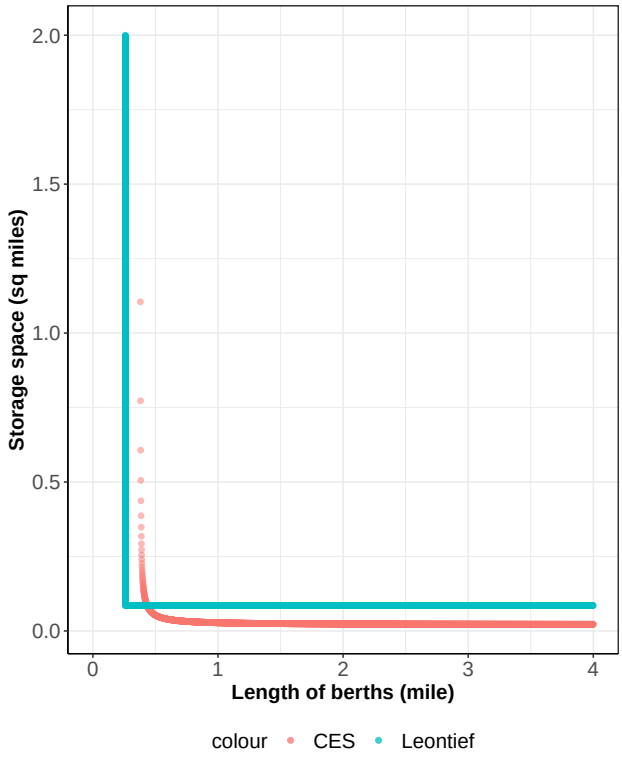


Figure 19: Capacity function. This figure compares the calibrated Leontief capacity function against an estimated CES capacity function and shows that the two are similar. In particular, we estimate $K_j = z_j \left[\alpha \left(\frac{B_j}{\kappa_1}\right)^{-\gamma} + (1-\alpha)\left(\frac{A_j}{\kappa_2}\right)^{-\gamma}\right]^{-\frac{1}{\gamma}}$ via GMM. We rewrite it as $\log K_j = \log \bar{z} + \epsilon_j - \frac{1}{\gamma}\left(\alpha \left(\frac{B_j}{\kappa_1}\right)^{-\gamma} + (1-\alpha)\left(\frac{A_j}{\kappa_2}\right)^{-\gamma}\right)$ and use the following two moment conditions: $E\left[B_j \epsilon_j\right] = 0$ and $E\left[A_j \epsilon_j\right] = 0$ to obtain the parameters.

	$\log(\omega_{jt})$
Tidal range (log)	0.155
	(0.057)
Port Type: Coastal natural	−0.191
	(0.0801)

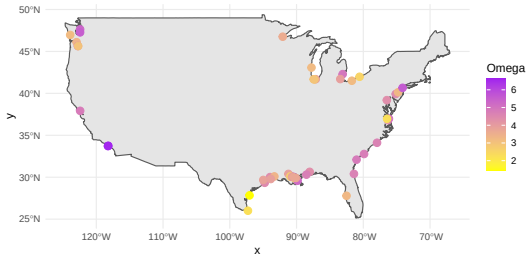


Table 7: Port Productivities. The table on the left presents the results from a regression of the estimated port productivity on tidal range and an indicator for port type obtained from the World Port Index (WPI). Coastal natural ports are sheltered from the wind by virtue of their location within a natural indentation, in comparison to coastal breakwater and river ports. The right panel plots the estimated average port productivities, $\log \omega_{jt} = \log \omega_j + \theta_\omega t + \nu_{jt}$, as obtained from the estimation of the service time CES specification (8).

	Cobb Douglas	CES
α_{cd}	0.88	
β_{cd}	0.60	
α_{ces}		0.76
η		-0.16
N	263	263

Table 8: This table reports the estimated parameters from the service time CES specification (8), as well as a Cobb-Douglas specification, where α_{cd} is the exponent on L_{jt}/s_{jt} and β_{cd} is the exponent on c_{jt}/s_{jt}.

Model:	(1)
Variables	
Unexpected disruptions to port operations	0.334***
	(0.070)
Foreign distance (days of travel)	-0.012
	(0.021)
Port price per ton	0.297
	(0.186)
Fixed-effects	
Port Region	Yes
Month	Yes
Foreign location	Yes
Fit statistics	
Observations	9,141
R^2	0.29
Wald (1st stage)	22.7

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

Table 9: "First stage". This table reports the regression of time at port on the instrument (unexpected service time disruptions).

	γ_t (1)	$\hat{\rho} y_t$ (2)
Constant	-0.0106	0.0479
	(0.9649)	(0.0282)
coal price	0.0056**	
	(0.0022)	
agricultural price index	0.0279**	
	(0.0125)	
base metal price	0.0035	
	(0.0026)	
$\hat{\rho} y_{t-1}$		0.9884
		(0.007341)
N	66	379
R^2	0.52	0.97

Table 10: This table reports the estimated parameters ρ that allow us to extrapolate the demand shock estimates γ_t to a longer time-series in the first column, as well as the estimated coefficients for the AR(1) process for γ_t given in (7) in the second column.

	Change in Wait Time (%)	Change in Trade (%)	Change in Welfare (%)	Rate of Return from Investment (%)
Alliance	-3.62	0.01	0.02	37.57
Ama	-22.13	0.31	0.73	440.68
Baltimore	-5.61	0.29	0.41	135.10
Baton Rouge	-6.83	0.02	0.06	36.86
Beaumont	-2.37	0.01	0.03	36.93
Brownsville	-2.44	0.01	0.01	19.37
Burnside	-24.86	0.08	0.20	197.56
Camden	-5.94	0.04	0.06	17.36
Charleston	-6.67	0.09	0.14	43.64
Cleveland	-0.08	0.00	0.00	0.02
Conneaut	-3.67	0.00	0.00	0.70
Convent	-8.85	0.20	0.49	725.07
Corpus Christi	-4.60	0.13	0.22	253.38
Darrow	-11.34	0.06	0.16	206.16
Davant	-0.50	0.00	0.00	7.25
Destrehan	-13.75	0.41	0.94	1074.82
Detroit	-9.32	0.01	0.01	6.48
Duluth	-0.06	0.00	0.00	0.15
East Chicago	-3.12	0.00	0.00	0.57
Fairless Hills incl. Morrisville	-4.88	0.03	0.04	15.39
Galveston	-9.17	0.15	0.28	110.51
Gramercy	-19.96	0.05	0.13	138.33
Grays Harbor	-12.49	0.19	0.28	170.60
Hampton Roads	-13.66	0.63	0.75	295.16
Houston	-1.61	0.10	0.21	148.47
Jacksonville	-7.26	0.04	0.07	31.90
Kalama	-9.65	0.77	0.99	769.63
La Place	-4.79	0.01	0.04	39.50
Lake Charles	-5.50	0.05	0.10	106.02
Long Beach	-12.22	0.36	0.48	91.67
Longview	-7.26	0.41	0.55	206.56
Los Angeles	-5.47	0.03	0.06	11.71
Milwaukee	-8.55	0.01	0.01	1.73
Mobile	-3.48	0.10	0.23	159.00
New Orleans	-4.57	0.16	0.39	137.93
Newark	-5.67	0.08	0.12	22.04
Pascagoula	-12.70	0.09	0.20	122.39
Port Arthur	-10.79	0.15	0.32	371.81
Portland	-6.98	0.40	0.51	65.09
Reserve	-3.47	0.05	0.14	154.52
Richmond (California)	-13.30	0.14	0.20	38.91
Savannah	-10.05	0.20	0.29	79.15
Seattle	-7.11	0.06	0.10	7.06
St Bernard Port	-8.66	0.04	0.10	71.51
Tacoma	-11.88	0.40	0.49	126.24
Tampa	-3.17	0.07	0.13	29.37
Toledo	-0.23	0.00	0.00	0.58
Vancouver	-12.45	0.47	0.62	179.53

Table 11: Welfare results. For each port in our sample, the table reports the change in average wait times, aggregate trade, and aggregate welfare, when expanding the port's capacity so it can handle one more vessel, as well as the rate of return from the investment (the cost is just recovered when the rate of return is equal to 100).

	Welfare Gains (log)	
δ_{geo} (standardized)	3.10	
	(2.05)	
Queue (standardized)	3.11	
	(0.46)	
δ_{geo} of substitute ports (standardized)		5.84
		(2.05)
Queue of substitute ports (standardized)		3.11
		(0.46)
R^2	0.75	0.55
N	49	49

Table 12: Determinants of welfare gains: This table regresses welfare gains on both the attributes of the treated port and those of its ten closest substitutes. We standardize the covariates to ease interpretation of the coefficients.

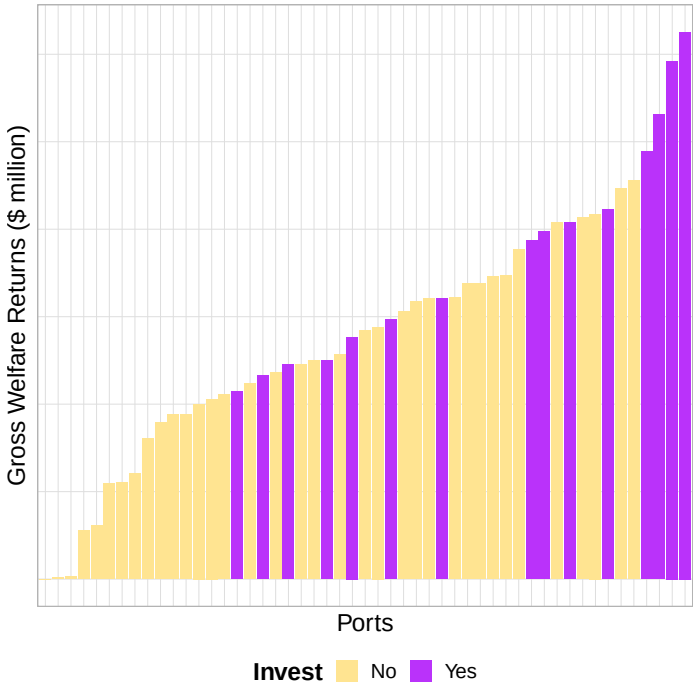


Figure 20: Welfare gains from investment in US ports. Ports are ranked according to their calculated welfare gains. The color varies depending on whether the gains are higher (purple) or lower (yellow) than the cost.

B Creation of Port Infrastructure Dataset

In this Appendix, we discuss the creation of the port infrastructure dataset.[53] Because ports seldom share information about their infrastructure, we use geolocation data and satellite imagery (provided by Google Maps and Google Earth Pro) to collect this information. Our approach consists of combining a number of different data sources (AXS Marine, Marine Traffic, Google Maps and Google Earth Pro), as well as a set of well-defined rules and extensive collaboration among data-collectors to achieve a reliable dataset on port infrastructure. We next describe the steps of our strategy. It is worth mentioning that ports are greatly heterogeneous (e.g. ocean vs. river) which makes each port unique in terms of data collection. This is also why current AI techniques may have difficulty collecting this information.

First, we obtain port polygons from AXS Marine. The polygons come in the form of Google Maps and highlight the outline of the port and its extent. This outline marks the geographic borders of the port. This is the area in which the data collectors need to identify different aspects of port infrastructure, including berths, storage, cranes, and silos. Figure 21 provides examples of polygons, which illustrate the considerable geographic heterogeneity. In particular, it depicts the polygons for Duluth (MN) and Baltimore (MD) which are visible different: Duluth is a simple harbor covering a relatively small area. In contrast, Baltimore is much more complex, as it includes both rivers and other coastal areas and is very extensive, so that a large area needs to be checked for port infrastructure.

We cross-validate this information by using public information from Marine Traffic, which also provides locations for some of the berths. We add any missing terminals/berths to the port. In addition, Marine Traffic contains information on the type of berth: we only want to include berths that handle cargo (general cargo, dry bulk, but also oil and containers) and thus remove berths that cater to fishing ships, cruise ships, military ships, and recreational vessels. We designate each berth to its purpose (dry bulk, oil, or containers).

In the next step, the data collectors measure several features of port infrastructure for every port: the length of every berth, the storage areas and silos, the number of cranes, and whether the port has railroad access, and/or a refinery nearby. Figure 22 provides one illustrative example from the port of Newark; on the image one can see clearly a ship parked at berth, the designated storage space next to it, as well as the cranes (un)loading the vessel. We next discuss each feature in turn.

Once all berths of a port are annotated in Google Maps, the data collectors locate them in Google

[53]We are most grateful to Pedro Degiovanni, as well as a Ian Bourgin, Victoria Healey, Shai Hirschl, Virginia Jablonski, Ellie Kominiarek, Christos Papageorgiou, Emily Taylor, Robert Yampanis, Jiayi Zhang for their careful and dedicated work in creating this dataset.

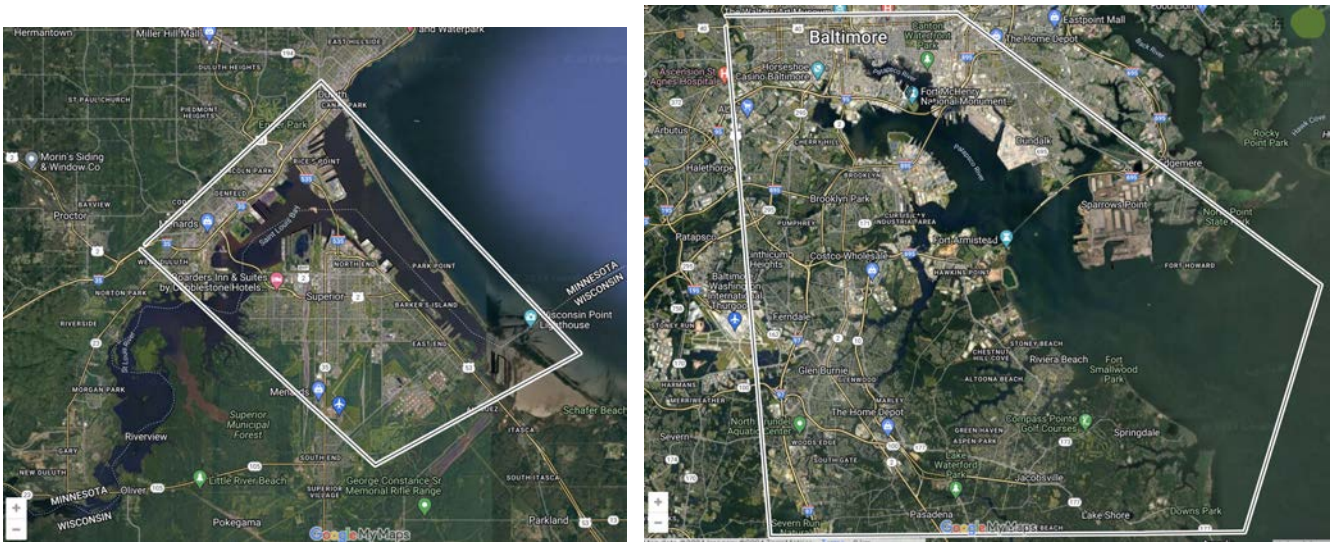


Figure 21: Port polygons for Duluth (MN) on the left and Baltimore (MD) on the right. These polygons capture the possible area of the port and are searched for port infrastructure, including berths, storage space, and cranes.

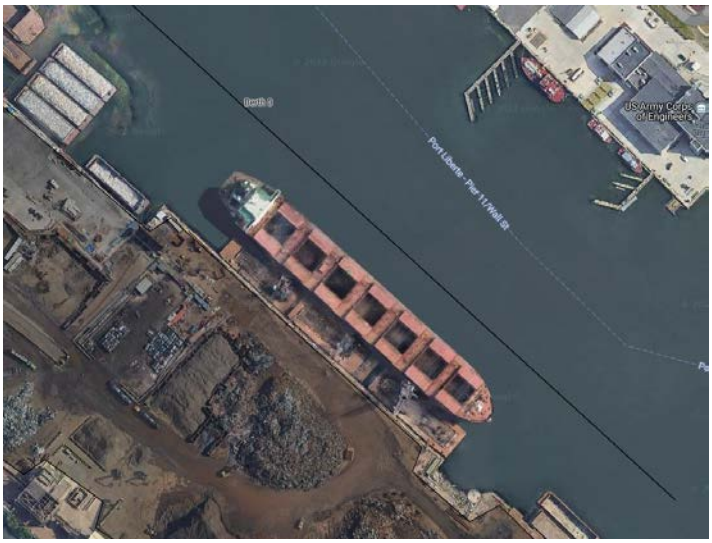


Figure 22: Satellite imagery of the port of Newark from Google Earth Pro. Note a ship parked at berth, the designated storage space next to it, as well as the cranes (un)loading the vessel.

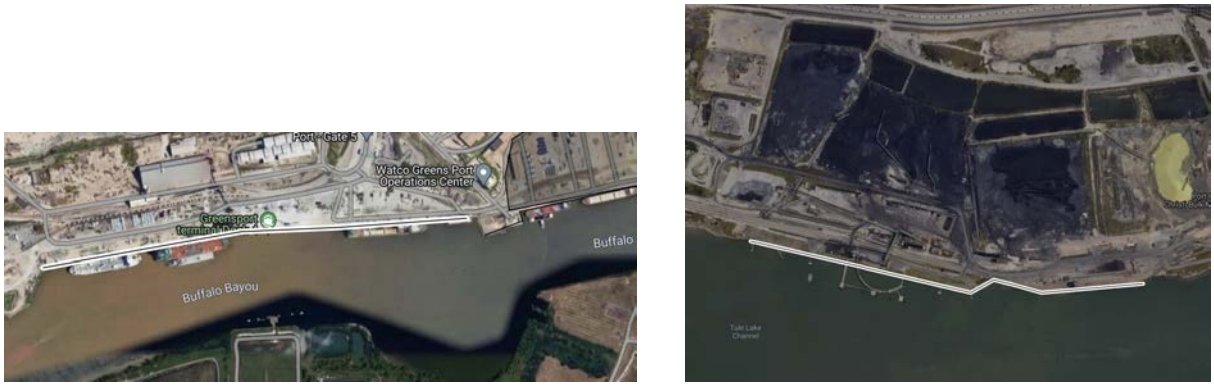


Figure 23: Different berths at the port of Houston (TX) on the left and Corpus Christi (TX) on the right.

Earth Pro. For each berth, they delineate its length and record it. Berths may have peculiar shapes so attention must be paid that all relevant berth space is included and that the measurement is done properly. As berths can have varying length, this measure of infrastructure is much more granular and informative than simply the number of berths. Figure 23 depicts examples of different berths from the ports of Houston (TX) and Corpus Christi (TX).

Ports have massive yards where they store cargo, such as bulk commodities, or containers. For each port, the data collectors first identify the different areas allocated to storage and for each of them they create a polygon on Google Maps. Then, they record its area in acres and indicate whether the storage is used for containers or not.

Figure 24: Different storage areas at Houston (TX) on the left and Hampton Roads (VA) on the right.

Encoding storage areas is likely the trickiest step of our approach and does in some cases require judgement calls by the data collectors. In those cases, they were asked to cross-check their work with their co-workers and mark their assessments in the dataset. For instance, coal is generally easy to recognize

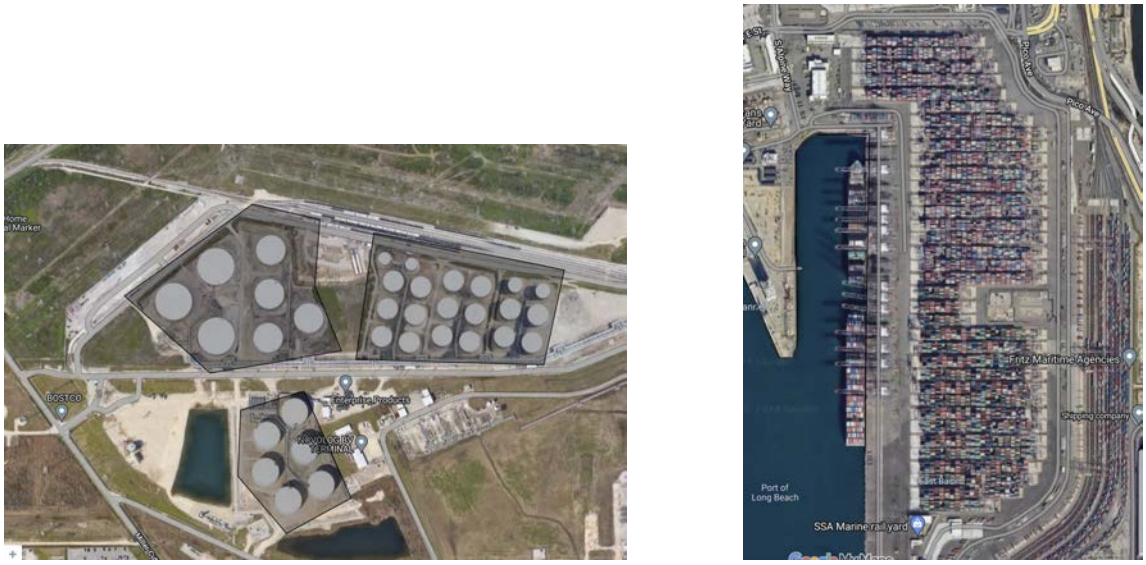


Figure 25: The left picture depicts silos areas at Houston (TX). The right picture depicts container storage space in Long Beach (CA).

due to its color, and so are piles of metal, but other minerals may appear more like dirt (this is less of an issue in the US). Similarly, some storage sites may initially look like construction, but they are in fact deposits of minerals. Container storage is easier to identify. However, while some ports have large, distinct areas of container storage, other ports have more small, scattered areas, or areas that look like junkyards. Another issue involves distinguishing between storage of cars and parking lots.[54] To address the potential for measurement error when it comes to storage areas, we include in our analysis only storage areas that are within a one mile radius from a berth, and perform robustness by using storage areas within a three mile radius. Figure 24 depicts examples of storage areas in Houston (TX) and Hampton Roads (VA). The right picture in Figure 25 shows an example of container storage from Long Beach (CA).

Silos are often used to store grain, oil or gas and are an important part of storage for commodities. Data collectors count silos, which are giant cylinders that look like circles on Google Earth Pro and report their total area as part of storage.[55] The left picture in Figure 25 depicts silos in Houston (TX).

Next, we measure the number of cranes. It is easier to see cranes in sunnier places due to the shadows they cast; so oftentimes the data collectors have to search for a good recent satellite image on a sunny day to spot them. Note for instance the cranes in the port of Newark in Figure 22.

We also track whether the port is connected to a railroad or not, as well as whether there is a refinery nearby. In the US, all ports in our sample are connected to railroad.

[54] In this case we adopted the following guidelines: if we see containers with a truck attached to them, we consider it a parking lot. If we see traffic signs and driving space between cars, we consider it a parking lot. If the cars are all packed together we count it as storage.

[55] To distinguish silos from other shorter cylindrical objects, we check whether they cast a shadow.

Finally, we use Google Earth Pro (in particular the "historical imagery" function) to construct a time series for all the features collected, from 2010 until 2021.

C Cost of Port Infrastructure

In this Appendix we provide all details for the construction of the cost of infrastructure from Section 4.3. We begin with the cost of dredging and then move to the cost of land.[56]

Dredging Costs As discussed in Section 4.3, there are two main steps to estimating dredging costs. For each port j, we need (i) a port- or berth- level measure of the seafloor depth at which rock material begins, denoted d_j^r; (ii) an estimate of the function linking dredging costs, to the quantity of rock and non-rock material extracted. Once the function from (ii) is known, we use the amount of rock and non-rock material that needs to be extracted at each port to compute port-specific dredging costs. We describe each step in turn.

For the first step we rely on the National Crustal Model (NCM), currently in development by the US Geological Survey (USGS). This uses a large set of existing geophysical data – including measures of porosity, lithology, compressional and shear wave velocities, and depth to bedrock and basement material – to estimate an array of geophysical properties on a 1-km by 1-km grid spanning the entire contiguous United States. For each berth in our dataset, we use unpublished meter-by-meter estimates of subsurface compressional (i.e., primary) wave velocities estimated by the NCM that the USGS has provided. The model structure, calibration, and estimation procedure used to derive these estimates are developed in Boyd (2020).[57] From these meter-by-meter estimates, we calculate the depth (in meters) below each berth at port j at which primary wave velocity first exceeds 2,300 meters per second. This threshold is used as the point of delineation between sediment (i.e., non-rock) and rock material, whereafter drilling and blasting methods must be used to prepare subsurface material for dredging. The choice of 2,300 m/s primary wave velocity to delineate rock from non-rock material follows previous literature (see, e.g., MurphyIII, Ward, Boyd, IV, Nolen-Hoeksema, Art, and Rosales-R (2011) and sources therein) and is informed by measured ripping productivity of subsurface materials at different primary wave velocities.[58]

[56]The analysis in this section would not have been possible without the invaluable research assistance of Zach Saunders.

[57]We are most grateful to Oliver Boyd for making this data available to us.

[58]Generally, most rock types transition from 'rippable' to 'marginal' between 2,000 and 3,000 m/s, whereas the transition to 'non-rippable' generally occurs between 3,000 and 4,000 m/s. While these measures might suggest a threshold of 3,000 m/s be used as the point at which drilling and blasting become necessary for the preparation of rock for excavation, we note that measured ripping productivity (in cubic meters of ripped material per hour) declines precipitously well before rock is classified as either marginal or completely non-rippable. MurphyIII, Ward, Boyd, IV, Nolen-Hoeksema, Art, and Rosales-R

We now turn to the second step: estimating the cost of dredging as a function of the quantity of sediment that is excavated and depending on its consistency (rock vs. non-rock). Over the last two decades, the US Army Corps of Engineers (USACE) has conducted several large-scale dredging projects at major US harbors. For two such projects – the Delaware River Main Channel Deepening Project and the Boston Harbor Deep Draft Navigation Improvement Project, which broke ground in 2010 and 2017 respectively – the USACE provides detailed estimates of predicted dredging costs. For each proposed design depth under consideration, estimates were produced for the quantity of rock and non-rock material to be dredged, total dredging and disposal-related costs both for rock and for non-rock materials, total drilling and blasting-related costs required to prepare rock material for dredging and disposal, and the price level at which these dollar costs were calculated. These estimates were provided not just at the project level, but also for each terminal berth and for each major section of harbor channel. Altogether, there are 78 distinct port-section-depth combinations for which the USACE provides dredging cost and material quantity estimates.[59]

After separating these rock and non-rock dredging costs, all additional costs – including items such as "Miscellaneous Costs," "Construction Contingency," "Planning Engineering and Design," "Supervision & Administration," "Construction Management," "Resetting of Aids to Navigation," and "Real Estate" – are apportioned as rock- and non-rock-related costs in proportion to the ratio of rock to non-rock-related dredging construction costs calculated above. Finally, all prices are converted to 2022 USD by adjusting for inflation using the PPI.

We then regress linearly the USACE costs on the quantity of excavated material, in order to obtain the dredging cost functions, $C_j^{d,r}(q^r), C_j^{d,nr}(q^{nr})$. As independent variables we use the quantities of rock and non-rock material that need to be dredged for each project-segment-depth reported by the USACE, while the dependent variable is the associated cost estimated by the USACE as required to remove the quantity of the material. The fit for this specification is quite good, as shown in Figure 26.

Finally, we use this function to extrapolate to all ports and obtain a measure of dredging there. The

(2011) correlate measured ripping productivity of a D-10 tractor ripper against measured compressional wave velocity for a number of sandstone ripping projects and show that measured productivity can fall very close to 0 m^3/h at seismic velocities of less than 2,500 m/s, well below the 'marginal' classification threshold for sandstone rock. Similar results can be found in productivity ranges reported in the 49th edition of the Caterpillar Performance Handbook.

[59]For the Delaware River deepening project, all data come directly from the 2002 Delaware River Main Channel Deepening Project Comprehensive Economic Reanalysis Report, Appendix A: Cost Estimate. For the Boston Harbor deepening project, all cost and quantity estimates are retrieved from the 2013 Boston Harbor Navigation Improvement Study Final Feasibility Report, Appendix D2 – Project Cost Estimates.

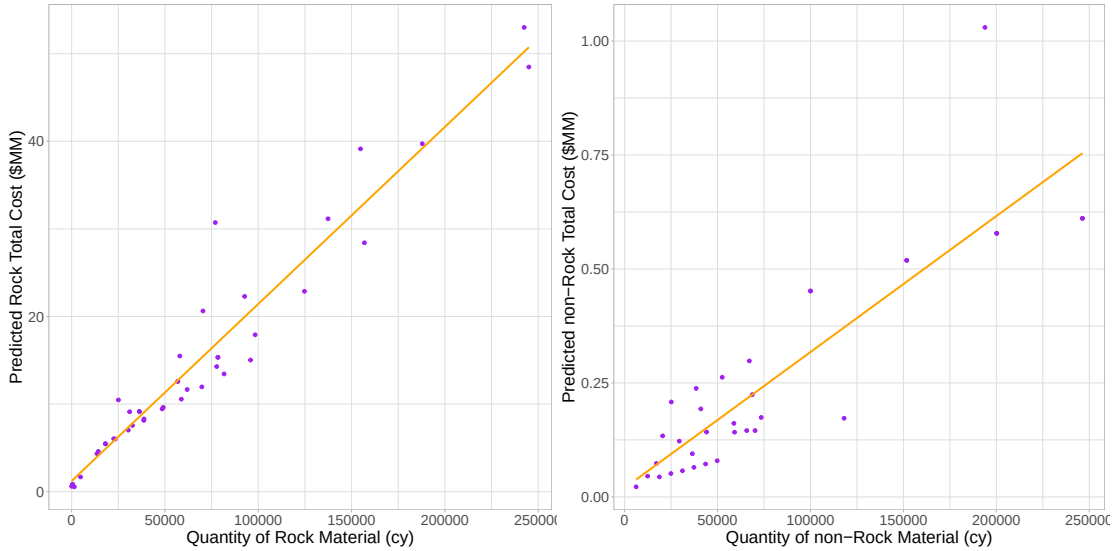


Figure 26: Dredging costs for rock and non-rock material. The dots correspond to different quantity-cost pairs for each project-berth-depth reported by the USACE and the solid line represents the estimated functional form of the dredging cost function.

total dredging cost is equal to the cost of dredging rock plus the cost of dredging non-rock material:

$$C_j^d(l, w, d) = C_j^{d,r}(q^r) + C_j^{d,nr}(q^{nr})$$

where q^r (q^{nr}) is the quantity of rock (non-rock) material. The task now is to compute q^r and q^{nr} for any port. We have

$$q^r = \mathbb{I}\left\{\left(d_j^r \leq d+2\right) \bigcap (d_j < d+4)\right\} lw\left((d+4) - \max\left\{d_j, d_j^r\right\}\right)$$

and

$$q^{nr} = \max\left\{0, lw\left(\min\left\{d_j^r, d+2\right\} - d_j\right)\right\}$$

where d_j is the current depth of port j, d_j^r is the depth at which rock begins at port j, d is the desired depth at which digging will take place, and l, w, d is the length, width and depth/draft of the average ship respectively. Let us explain these expressions:

First, suppose that $d_j \leq d+2 \leq d_j^r$, i.e., the new berth can be dredged to depth $d+2$ (to service ships of depth d with the customary 2 ft of additional overdepth) without hitting any non-dredgeable rock material. In this case, no rock material is dredged, $q^r = 0$, while the quantity of non-rock dredged material is given by $q^{nr} = lw\left((d+2) - d_j\right)$.

Second, suppose that $d_j^r \leq d_j \leq d+2$. In this case, all material to be excavated consists of non-

dredgeable rock material; i.e., $q^{nr} = 0$ and $q^r = lw\left((d+4) - d_j\right)$. Given that the bottom of the berth now consists of exposed rock, the customary 4 ft of overdepth must be provided to prevent damage to ships' hulls resulting from accidental contact with the rocky berth floor.

Third, suppose $d_j \leq d_j^r \leq d+2$. Then, non-rock material is dredged from the actual depth d_j to the depth of hitting rock d_j^r, while non-dredgeable rock material must be removed from depth d_j^r to depth $d+4$, in which case $q^n = lw\left(d_j^r - d_j\right)$ and $q^r = lw\left((d+4) - d_j^r\right)$.

For the purposes of this exercise, we set d_j equal to zero ft, as we consider this a brand new berth; we have however performed robustness with 5 ft and 10 ft, as a minimum depth below the water level; in these cases, costs are lower and a few more ports exhibit positive net returns. Moreover, the length of the berth is given by parameter κ_1 of the Leontief capacity function (9) as described in Section 4.1. We set the width (depth) equal to the width and draft of the average vessel.

Land Costs The second term of the cost function for investment in infrastructure consists of land costs required for additional storage. As discussed in Section 4.3, we estimate the cost of land acquisition by first using open-source map data to identify parcels of industrial and commercial land near existing port infrastructure, and then we use highly granular land value estimates from the ecology and conservation literature to assign unit land prices to these parcels.

In particular, we use OpenStreetMap (OSM) to extract all land-based features with either an industrial or commercial use located within 1, 2, or 3 miles of any berth affiliated with each port in our data.[60] The distance used for each port is the minimum distance sufficient to capture at least one parcel of industrial or commercial land neighboring a local terminal.[61]

We next calculate the monetary value of this industrial land that we have identified as local for each port. To do so, we rely on Nolte (2020) who provides land values for the US. In particular, Nolte (2020) uses data on land sales to train a predictive model for expected sale prices for all private properties in the contiguous United States.[62] The resulting property-level land price estimates are then rasterized into

[60] OSM is an open-access online map service updated and maintained by its users and comprises a highly-detailed database of natural and man-made geographic features. Among other information, OSM includes shape files representing the outline and location of each constituent geographic feature along with a land-use variable indicating the primary use of each land-based feature. The land-use variable takes 80 possible values, including residential, commercial, and industrial, among others.

[61] In two cases – the Port of Los Angeles and the Port of Long Beach – a single geographic feature must be dropped following this procedure, as it corresponds to a large area of Los Angeles Harbor that has been incorrectly coded as industrial land in OpenStreetMap.

[62] In particular, Nolte (2020) uses a tree-based bagging algorithm to estimate the expected sale price of all private land in the contiguous United States, with the unit of observation being the tax-assessor parcel. The model is trained on 6 million private land sales taking place between 2000 and 2019, and estimated land price is modeled as a function of publicly-available data on land ownership, sale information, building footprints, terrain, accessibility, land cover, hydrology, flood risk,

a 6,054 by 9,618 element raster covering the landmass of the contiguous US. Thus, each resulting raster element corresponds to an approximately square parcel of land with side length measuring 480 meters and is assigned a single value measuring its estimated price-per-hectare.

Using all such parcels of industrial and commercial land that are local to the ports in our dataset, we calculate the price-per-hectare of local industrial and commercial land as the weighted average of rasterized land value estimates from Nolte (2020), with weights corresponding to the amount of industrial and/or commercial land falling within each square of the land value grid. Because the land values are measured in 2010 USD, we then adjust for inflation using the PPI.

Finally, we compute the total land costs as the product of land prices and the acreage required to handle the new vessel, which is given by parameter κ_2 of the Leontief capacity function (9) as described in Section 4.1.

Bulkhead Costs A bulkhead is a vertical shoreline stabilization structure that primarily retains soil, and provides minimal protection from waves. As already mentioned, data on bulkhead construction are rare, however we are able to use reported information from the Charleston (SC) Harbor port expansion project. As part of this project, the construction of a new bulkhead was necessary, which – at the project planning stage – had an associated estimated cost of $22 million measured at April 2014 prices. The terminal has 3,800 linear ft of berthing space, which results in a rough cost estimate of $5,789 per linear ft of bulkhead. After adjusting for inflation, we let bulkhead costs amount to $6,600 per linear foot of berth in 2022 USD. This estimate is in line with one more reported bulkhead construction project we were able to find, the 1,671-ft Port Fourchon bulkhead construction project completed by the Greater Lafourche Port Commission in 2019, which cost approximately $11 million, corresponding to a cost-per-foot of roughly $6,582 in 2022 USD.

D Counterfactual Simulations

In this Appendix, we describe the algorithm employed to compute the equilibrium of our model. Every month t, the importer-exporter pairs face an aggregate demand state γ_t, while every port j offers a service

demographics, and environmental protections. The trained model and underlying data are then used to predict expected sale prices for all private properties in the contiguous United States, expressed in 2010 USD. These estimates are validated against verified sale prices of 4,128 publicly funded land acquisitions in 659 US counties, and found to explain approximately 70% of variation in log per-area sale prices.

rate $\mu_{jt}(\equiv 1/T_{jt})$ and has capacity K_j.[63]

Finding the equilibrium of our model requires computing, at every port $j \in J$ and month $t \in \{1,\ldots,T\}$, the following three objects. First, the probabilities $\{p_{n,jt}(\tau_h)\}_{n,h}$ that there are n ships at port j at instant τ_h of month t, for a grid $\{\tau_1,\ldots,\tau_H\}$ of half-hour intervals within the month. Second, the average total time at port TT_{jt} implied by these probabilities. Third, the arrival rate λ_{jt} of ships at port j. Note that, since traders make their decisions based on the expected time at port TT_{jt}, the arrival rate λ_{jt} will depend on this object. In turn, the state of the queue at each port, described by the probabilities $\{p_{n,jt}(\tau_h)\}_{n,h}$ depends on the ships' arrival rates λ_{jt}. Our definition of equilibrium requires that the arrival rate λ_{jt} is consistent with TT_{jt} given our demand model, and that the probabilities $\{p_{n,jt}(\tau_h)\}_{n,h}$ are consistent with the arrival rate λ_{jt}, given the differential equations (3) describing the evolution of the queue. Also note that in every month the initial state of the queue depends on the residual queue at port at midnight on the last day of the previous month, which is described by the probabilities $\{p_{n,jt}(\tau_H)\}_n$. Beyond this "initial condition" the evolution of the queue is independent across months. We next describe how we compute these equilibrium objects within each month.

The algorithm proceeds as follows:

1. Make an initial guess for the expected time at port in month t, $TT_t^0 = \left(TT_{1,t}^0,\ldots,TT_{J,t}^0\right)$.

2. At each iteration k, we inherit TT_t^{k-1}. We then update our guess according to the following steps.

   (a) Compute the number of ships arriving at each port, $n_t^k = \left(n_{1t}^k,\ldots,n_{Jt}^k\right)$, given the expected time at port TT_t^{k-1}, from

   $$n_{jt}^k = \sum_{fd} \frac{\exp\left(\delta_{jft}^k + \beta_D \text{dist}(j,d)\right)}{1+\sum_l \exp\left(\delta_{lft}^k + \beta_D \text{dist}(l,d)\right)} m_{fd} M_f,$$

   where

   $$\delta_{jft}^k = \beta_F \text{dist}(f,j) - \beta_T TT_{jt}^{k-1} - \beta_P p_j + \gamma_t + \gamma_f + \gamma_{l(j)} + \xi_{jft}. \tag{12}$$

   Denote the corresponding arrival rates by $\lambda_{jt}^k = n_{jt}^k$.

   (b) Next, compute the evolution of the queue in each port *within* month t, given the arrival rates λ_{jt}^k, service rates μ_{jt}, capacities K_j, and probabilities $p_{n,jt-1}(\tau_H)$. To this end, we consider a

[63] In our simulations we set the service rate $\mu_{jt} = 1/T_{jt}$, where T_{jt} is the average service time in port j and month t. In practice, we use average service time over a six month period. Note that, as discussed in Section 4.3, in the simulations the service rate remains fixed at μ_{jt} when the port capacity increases.

grid $\{\tau_1, \ldots, \tau_H\}$ of half-hour intervals within the month and apply the differential equations (3) to each port j as follows:

i. For half-hour $h = 1$ we approximate the probability $p^k_{n,jt}(\tau_1)$ that there are n ships at port j at instant τ_1 of month t using

$$p^k_{n,jt}(\tau_1) = p^k_{n,j,t-1}(\tau_H) + (\tau_h - \tau_{h-1}) \frac{d}{d\tau} p^k_{n,jt-1}(\tau_H)$$

where

$$\frac{d}{d\tau} p^k_{n,jt-1}(\tau_H) = \begin{cases} -\lambda^k_{jt} p^k_{0,jt-1}(\tau_H) + \mu_{jt} p^k_{1,jt-1}(\tau_H) & \text{if } n = 0 \\ -\left(\lambda^k_{jt} + n\mu_{jt}\right) p^k_{n,jt-1}(\tau_H) + \lambda^k_{jt} p^k_{n-1,jt-1}(\tau_H) + (n+1)\mu_{jt} p^k_{n+1,jt-1}(\tau_H), & \text{if } 0 < n < K_j \\ -\left(\lambda^k_{jt} + K_j\mu_{jt}\right) p^k_{n,jt-1}(\tau_H) + \lambda^k_{jt} p^k_{n-1,jt-1}(\tau_H) + K_j\mu_{jt} p^k_{n+1,jt-1}(\tau_H), & \text{if } n \geq K_j \end{cases} \tag{13}$$

ii. Similarly, for each interval $h > 1$, we approximate

$$p^k_{n,jt}(\tau_h) = p^k_{n,jt}(\tau_{h-1}) + (\tau_h - \tau_{h-1}) \frac{d}{d\tau} p^k_{n,jt}(\tau_{h-1}),$$

where $\frac{d}{d\tau} p^k_{n,jt}(\tau_{h-1})$ is defined as above.

(c) Given the probabilities $\left\{p^k_{n,jt}(\tau_h)\right\}_{h=1}^H$ we can compute the expected time at port as

$$TT^k_{jt} = \frac{1}{H} \sum_h \sum_n \left(\frac{1}{\mu_{jt}} + \frac{(n - K_j + 1)}{\mu_{jt} K_j} \mathbb{I}\{n \geq K_j\} \right) p^k_{n,jt}(\tau_h),$$

where we use the probabilities $p^k_{n,jt}(\tau_h)$ to take expectations across the possible states of the queue at port j, and then take the average across the intervals τ_h.

(d) If $\max_j \left\{ \left| TT^k_{jt} - TT^{k-1}_{jt} \right| \right\} < \epsilon$, then we set $TT_{jt} = TT^k_{jt}$ and $\lambda_{jt} = \lambda^k_{jt}$.

In practice, in our counterfactuals we compute the equilibrium across 300 simulations, where we re-draw the monthly demand state γ_t. For each simulation m we set γ^m_0 equal to the estimated demand state on January 2018. Then, we simulate the demand state for thirty years according to

$$\gamma^m_t = \rho_0 + \rho_1 \gamma^m_{t-1} + \epsilon^m_t$$
$$\epsilon^m_t \sim N\left(0, \sigma^2_\epsilon\right),$$

where $(\rho_0, \rho_1, \sigma_\epsilon)$ are the estimated parameter of the AR process described in Section 4.2.

Finally, we compute expected discounted welfare by taking the average over the simulations. We compute the expected discounted welfare for each simulation as

$$\sum_t \beta^t W\left(TT_t, \gamma_t\right),$$

where

$$W\left(TT_t, \gamma_t\right) = \frac{1}{\beta_P}\left[\sum_{d,f} \log\left(\sum_j \exp\left(\beta_T TT_{jt} + \gamma_t + x_{jfd}\right) + 1\right) m_{fd} M_f + \gamma^{euler}\right]$$

i.e. the present discounted sum of consumer welfare across all options, $W\left(TT_t, \gamma_t\right)$, which depends on the vector of time at port TT_t, the aggregate demand γ_t, as well as x_{jfd} which includes the remaining port characteristics, such as the distances and prices. Given that port prices in our data vary little over time, even during the period when demand surged, we set them equal to their port-specific sample average. Similarly, we keep ξ_{jft} constant and equal to the average over time in our sample. We calibrate the discount factor, β, to 0.95 annually.

The algorithm is relatively fast and robust. The key to its speed is the use of equations (3), which governs the evolution of the queueing system. This approach, which is standard in queuing theory, allows us to not simulate the queueing process directly (i.e. tracking individual ship arrivals and service times) which would be substantially more computationally intensive.

Although we do not have a formal proof of equilibrium uniqueness, we have simulated our setup using different initial conditions and every time the algorithm leads to same equilibrium. In addition, unlike setups where there are strong strategic complementarities in agents' actions (often necessary to obtain multiple equilibria), here shippers' decisions are strategic substitutes: if many shippers choose a particular port j, then this lowers rather than increases one's incentives to pick j. Thus it would not be surprising if the equilibrium set of TT_js is unique.

References

Alessandria, G., J. P. Kaboski, and V. Midrigan (2010): "Inventories, lumpy trade, and large devaluations," *American Economic Review*, 100, 2304–2339.

Alessandria, G., S. Y. Khan, A. Khederlarian, C. Mix, and K. J. Ruhl (2023): "The Aggregage Effects of Global and Local Supply Chain Disruptions: 2020-2022," *Journal of International Economics*, 146, 103788.

Allen, T. and C. Arkolakis (2022): "The Welfare Effects of Transportation Infrastructure Improvements," *The Review of Economic Studies*, 89, 2911–2957.

Almagro, M., F. Barbieri, J. C. Castillo, N. G. Hickok, and T. Salz (2024): "Optimal Urban Transportation Policy: Evidence from Chicago," Tech. rep., National Bureau of Economic Research.

Anderson, J. E. and E. Van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–192.

Arellano, M. and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, 58, 277–297.

Arnott, R., A. De Palma, and R. Lindsey (1990): "Economics of a bottleneck," *Journal of Urban Economics*, 27, 111–130.

Asturias, J. (2020): "Endogenous Transportation Costs," *European Economic Review*, 123, 103366.

Bai, X., J. Fernández-Villaverde, Y. Li, and F. Zanetti (2024): "The causal effects of global supply chain disruptions on macroeconomic outcomes: evidence and theory," *NBER Working Paper*, 32098.

Bailey, S. (2021): "Competition and Coordination in Infrastructure: Port Authorities' Response to the Panama Canal Expansion," *Unpublished Manuscript*.

Berry, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

——— (2004): "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of Political Economy*, 112, 68–105.

Bhat, U. N. (2008): *An Introduction to Queueing Theory: Modeling and Analysis in Applications*, vol. 36, Springer.

Björkegren, D., A. Duhaut, G. Nagpal, and N. Tsivanidis (2025): "Public and Private Transit: Evidence from Lagos," Tech. rep., National Bureau of Economic Research.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Bonadio, B. (2024): "Ports vs. roads: infrastructure, market access and regional outcomes," *CESifo Working Paper*, 11383.

Bordeu, O. (2023): "Commuting Infrastructure in Fragmented Cities," *Job Market Paper, University of Chicago Booth School of Business*.

Boyd, O. (2020): "Calibration of the U.S. Geological Survey National Crustal Model," Tech. rep., US Geological Survey, USA.

Brancaccio, G., M. Kalouptsidi, and T. Papageorgiou (2020): "Geography, Transportation, and Endogenous Trade Costs," *Econometrica*, 88, 657–691.

——— (2025a): "Rigidities in Transportation and Supply Chain Disruptions," *AEA Papers and Proceedings*, 115, 543–550.

Brancaccio, G., M. Kalouptsidi, T. Papageorgiou, and N. Rosaia (2023): "Search Frictions and Efficiency in Decentralized Transport Markets," *Quarterly Journal of Economics*, 138, 2451–2503.

Brancaccio, G., M. Kalouptsidi, T. Papageorgiou, and Y. Zhou (2025b): "Port Disruptions and Prices," *mimeo.*

Brooks, L., N. Gendron-Carrier, and G. Rua (2021): "The Local Impact of Containerization," *Journal of Urban Economics*, 126.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi (2021): "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake," *The Quarterly Journal of Economics*, 136, 1255–1321.

Conlon, C. and J. Gortmaker (2020): "Best practices for differentiated products demand estimation with pyblp," *The RAND Journal of Economics*, 51, 1108–1161.

Conwell, L. (2024): "Privatized Provision of Public Transit," *mimeo, University College London.*

Cook, C. and P. Z. Li (2023): "Value pricing or lexus lanes? the distributional effects of dynamic tolling," Tech. rep., mimeo, Stanford University.

Donaldson, D. (2012): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *forthcoming, American Economic Review.*

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

Downs, A. (1962): "The Law of Peak-Hour Expressway Congestion," *Traffic Quarterly*, 16, 393–409.

——— (1992): *Stuck in traffic: Coping with peak-hour traffic congestion*, Brookings Institution Press.

Ducruet, C., R. Juhasz, D. K. Nagy, and C. Steinwender (2024): "All Aboard: The Effects of Port Development," *Journal of International Economics*, 151, 103963.

Duranton, G. and M. A. Turner (2011): "The fundamental law of road congestion: Evidence from US cities," *American Economic Review*, 101, 2616–2652.

Durrmeyer, I. and N. Martinez (2022): "The welfare consequences of urban traffic regulations," *TSE Working Paper*.

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

Elliott, M., B. Golub, and M. V. Leduc (2022): "Supply network formation and fragility," *American Economic Review*, 112, 2701–2747.

Fajgelbaum, P. D. and E. Schaal (2020): "Optimal Transport Networks in Spatial Equilibrium," *Econometrica*, 88, 1411–1452.

Friedt, F. L. and W. W. Wilson (2020): "The Economic Impact of Seaport and Other Infrastructure Investments and Leakages: A Literature Review," *Journal of Transport Economics and Policy*, 25, 219–243.

Fuchs, S. and W. F. Wong (2025): "Multimodal Transport Networks," *mimeo, University of Oregon*.

Ganapati, S., W. F. Wong, and O. Ziv (2024): "Entrepot: Hubs, Scale, and Trade Costs," *American Economic Journal: Macroeconomics*, 16, 239â78.

Glaeser, E. L. and J. M. Poterba (2021a): *Economic Analysis and Infrastructure Investment*, Chicago: University of Chicago Press.

——— (2021b): "Economic Perspectives on Infrastructure Investment," in *Rebuilding the Post-Pandemic Economy*, ed. by M. S. Kearney and A. Ganz, Aspen Institute Press.

Grossman, G. M., E. Helpman, and H. Lhuillier (2021): "Supply Chain Resilience: Should Policy Promote Diversification or Reshoring?" *NBER Working Paper*, 29330.

Heiland, I., A. Moxnes, K. H. Ulltveit-Moe, and Y. Zi (2025): *The Review of Economics and Statistics*, 1–45.

Hierons, T. R. (2025): "Optimal Congestion Pricing in General Equilibrium," *mimeo, University of Chicago*.

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

Hummels, D., V. Lugovskyy, and S. Skiba (2009): "The Trade Reducing Effects of Market Power in International Shipping," *Journal of Development Economics*, 89, 84–97.

Hummels, D. and S. Skiba (2004): "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 112, 1384–1402.

Ishikawa, J. and N. Tarui (2018): "Backfiring with Backhaul Problems: Trade and Industrial Policies with Endogenous Transport Costs," *Journal of International Economics*, 111, 81–98.

Kalouptsidi, M. (2014): "Time to Build and Fluctuations in Bulk Shipping," *American Economic Review*, 104, 564–608.

Kreindler, G. (2024): "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications," *Econometrica*, 92, 1233–1268.

Larson, R. C. and A. R. Odoni (2007): *Urban Operations Research*, Dynamic Ideas, 2nd ed.

Lee, H. L. and M. A. Cohen (1983): "A note on the convexity of performance measures of M/M/c queueing systems," *Journal of Applied Probability*, 20, 920–923.

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

MurphyIII, W., W. Ward, B. Boyd, W. M. IV, R. Nolen-Hoeksema, M. Art, and D. Rosales-R (2011): "Geophysical, Geological, Geotechnical, and Mechanical Testing of Rock." *Proceedings of the Western Dredging Association (WEDA XXXI) Technical Conference and Texas AM University (TAMU 42) Dredging Seminar*, 423–435.

Nolte, C. (2020): "High-resolution land value maps reveal underestimation of conservation costs in the United States," *Proceedings of the National Academy of Sciences*, 117, 29577–29583.

Notteboom, T., A. Pallis, and J.-P. Rodrigue (2022): *Port economics, management and policy*, Routledge.

Redding, S. J. (2016): "Goods trade, factor mobility and welfare," *Journal of International Economics*, 101, 148–167.

——— (2025): "Evaluating transport improvements in spatial equilibrium," Tech. rep., National Bureau of Economic Research.

Redding, S. J. and M. A. Turner (2015): "Chapter 20 - Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 1339–1398.

Sytsma, T. and W. W. Wilson (2020): "Port Choice and International Trade in Agricultural Products," *mimeo, University of Oregon*.

Vickrey, W. S. (1969): "Congestion Theory and Transport Investment," *The American Economic Review*, 59, 251–260.

Wong, W. F. (2022): "The Round Trip Effect: Endogenous Transport Costs and International Trade," *American Economic Journal: Applied Economics*, 14, 127–166.