



**HAL**  
open science

# The effect of LNG bunkering on port competitiveness using multilevel data analysis

Akoh Fabien Yao, Maxime Sèbe, Laura Recuero Virto, Abdelhak Nassiri, Hervé Dumez

## ► To cite this version:

Akoh Fabien Yao, Maxime Sèbe, Laura Recuero Virto, Abdelhak Nassiri, Hervé Dumez. The effect of LNG bunkering on port competitiveness using multilevel data analysis. *Transportation Research Part D: Transport and Environment*, 2024, 132, pp.104240. <10.1016/j.trd.2024.104240>. <hal-04611804>

**HAL Id: hal-04611804**

**<https://hal.science/hal-04611804v1>**

Submitted on 14 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Transportation Research Part D

journal homepage: [www.elsevier.com/locate/trd](http://www.elsevier.com/locate/trd)

# The effect of LNG bunkering on port competitiveness using multilevel data analysis

Akoh Fabien Yao<sup>a,\*</sup>, Maxime Sèbe<sup>a</sup>, Laura Recuero Virto<sup>b</sup>, Abdelhak Nassiri<sup>c</sup>, Hervé Dumez<sup>a</sup>

<sup>a</sup> i3-CRG, UMR CNRS 9217, École polytechnique, Institut polytechnique de Paris, 828 Boulevard des Maréchaux, 91120 Palaiseau, France

<sup>b</sup> Leonard de Vinci Pole Universitaire, Research Center, 92916 Paris La Defense, France

<sup>c</sup> UMR 6308 Amure, Université de Bretagne Occidentale, France

## ARTICLE INFO

### Keywords:

LNG  
Port choice  
Baltic Sea  
Multilevel regression  
Propensity score matching

## ABSTRACT

Environmental practices can enable most businesses, including ports, to gain a competitive advantage. Given the chicken-and-egg dilemma for the adoption of alternative fuels in the shipping industry, this article assesses whether and to what extent ports have incentives to provide liquefied natural gas (LNG) bunkering infrastructure. More specifically, we test whether such facilities positively affect the competitiveness of the ports, which would be an additional incentive to drive the transition to alternative fuels. Using multilevel regressions and propensity score matching on LNG-fueled vessel movements in the Baltic Sea Region in 2019, we found no significant effect of LNG bunkering infrastructure on port competitiveness, measured by port choice probabilities expressed by vessels. Although our findings indicate that ports do not gain a competitive advantage in the short-term, we do not rule out potential gains in the long-term. Policy intervention is desirable in the short-term to maintain incentives for port investments.

## 1. Introduction

Maritime transportation is the most energy-efficient means of transporting goods. Nonetheless, it already generates 3 % of global anthropic greenhouse gas (GHG) emissions (Fan et al., 2020; IMO, 2015). If shipping were a country, it would be the 6th largest producer of GHG (Eide et al., 2009). It is the most common means of transportation for international trade, with 90 % of traded goods carried by sea (UNCTAD, 2020). Furthermore, due to the growth in international trade, the impact of shipping is expected to reach 17 % of total GHG emissions by 2050 (ITF - OECD, 2018). Alternative fuels, such as liquefied natural gas (LNG) or hydrogen, are expected to emerge for the mitigation of shipping emissions. Indeed, the use of LNG (Acciaro, 2014) and hydrogen reduces the emissions of sulfur oxide (SO<sub>x</sub>) to zero versus conventional marine gasoil. Nitrogen oxide (NO<sub>x</sub>) emissions are also reduced by 90 % and 100 %, respectively (McKinlay et al., 2021; Wang and Notteboom, 2014).

The wide adoption of these fuels by shipping companies relies on the availability of bunkering infrastructure. At the same time, port authorities and bunkering operators will not provide such infrastructure until sufficient demand is expressed, leading to the so called ‘chicken-and-egg’ problem (Aronietis et al., 2016). Fostering the penetration of alternative fuels in this case requires policy interventions (Yeh, 2007). Most research on the subject has been devoted to identifying which of the ports or shipping companies should

\* Corresponding author.

E-mail addresses: [fabien.yao@polytechnique.edu](mailto:fabien.yao@polytechnique.edu), [fabienyao19@gmail.com](mailto:fabienyao19@gmail.com) (A.F. Yao).

<https://doi.org/10.1016/j.trd.2024.104240>

Received 21 December 2023; Received in revised form 21 March 2024; Accepted 8 May 2024

Available online 17 May 2024

1361-9209/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

be the focus of such policies (Zink et al., 2020), through e.g., subsidies for infrastructure investments or fuel adoption for vessels. However, evidence is lacking on the economic incentives of being a first mover in such a chicken-and-egg situation. As bunkering infrastructure for alternative fuels is considered to be a component of the environmental strategy of ports, a number of authors argue that it would increase their attractiveness (Kim and Chiang, 2014; Peng et al., 2021; Wang and Notteboom, 2015). However, there is no empirical evidence of this. The present article contributes to the literature by investigating the impact of alternative fuel (LNG) infrastructure on port competitiveness, measured as port choice probability. By analyzing the movements of LNG-fueled vessels in the Baltic Sea Region (BSR) in 2019, we model port choice probabilities and estimate the effect of providing LNG bunkering infrastructure, based on multilevel regressions and propensity score matching. The article hence attempts to highlight the potential market incentives for ports in providing bunkering facilities, given the low level of adoption of alternative fuels. The focus is on LNG because it was the most prevalent alternative fuel in the BSR in 2019 despite its small market share (1 %). Most vessels were powered by low-Sulphur fuel oil (LSFO, 74 %), marine diesel oil (MDO, 15 %), or heavy fuel oil (HFO) associated with an exhaust gas cleaning system (10 %). While there was no record of a hydrogen vessel to our knowledge, only one vessel was powered by methanol, making LNG the ideal choice for such data analysis.

The remainder of the article is organized as follows. A review of the existing literature on alternative fuel infrastructure and port competitiveness is presented in section 2. In the same section, we also highlight the role of ports in solving the chicken-and-egg problem and emphasize the expected gains in competitiveness attributed to such infrastructure based on port choice theory. Our data and methodology are presented in the third section and the results and concluding remarks are developed in sections 4 and 5, respectively.

## 2. Literature review

### 2.1. The role of ports in the adoption of maritime alternative fuels: A chicken-and-egg perspective

Infrastructure-dependent technologies tend to present the so-called “chicken-and-egg” problem, characterized by hesitation in the decision to invest by both the supply and demand sides. In the case of maritime alternative fuels, on the one hand, shipping companies are reluctant to adopt them as long as the supply and bunkering are uncertain and the costs remain high (Li et al., 2013). On the other hand, ports and bunkering operators will not provide the required infrastructure unless ship-owners express a significant demand (Aronietis et al., 2016; Wang and Notteboom, 2015). As a result, such hesitation maintains the market penetration of alternative fuels at a low level and the conundrum persists. The chicken-and-egg problem is not exclusive to the shipping industry. It arises during the deployment of most infrastructure-dependent technologies, with battery electric vehicles (BEV) being one of the most popular cases in the automotive industry. The degree of adoption of BEV by users depends on the availability of charging infrastructure. At the same time, the deployment of such infrastructure depends on the degree of adoption of BEV. In the presence of ongoing coordination failures between the two sides of the market, government intervention appears to be justified (Salmenkaita and Salo, 2002) and existing research has tried to provide theoretical and empirical guidance for policies.

Overcoming the chicken-and-egg problem is not straightforward. While most studies show that policy intervention should be targeted toward infrastructure provision (Delacrétaz et al., 2020; Mersky et al., 2016; Schulz and Rode, 2022), others have expressed a number of reservations. Considering several agents (policy makers, firms and consumers), Brozynski and Leibowicz (2022) found that although subsidies are most often optimal for infrastructure provision, this is not the case when the marginal social benefits<sup>1</sup> of the technology increase. In such a situation, subsidies should support adoption by consumers.

Concerning maritime transportation, alternative fuels have not yet fully emerged. They are still at the beginning of the adoption process compared to road transportation and BEV, and the role of bunkering infrastructure is no less crucial. Mäkitie et al. (2022) investigated the motivations and barriers for the adoption of alternative fuels by ship-owning companies in Norway, and found that the lack of infrastructure and the limited number of suppliers are substantial barriers common to all types of firms.<sup>2</sup> Alamouh et al. (2022) emphasized the role of ports as incentive providers, not only for the adoption of alternative fuels, but also for other measures of emissions reduction. Wang and Notteboom (2015) argued that the provision of infrastructure for alternative shipping fuels falls under the remit of the port authorities. They should embrace this role as part of their sustainability strategy and establish themselves as key players in their local development. As compared to a situation where ship-owners would invest in individual bunkering solutions, it would be collectively more advantageous that ports, as central nodes in the shipping network, provide such facilities. Besides, equipping vessels with alternative fuels without a viable bunkering solution may result in a paralysis of the whole supply chain.

Despite being the most adopted alternative fuel in the shipping industry, LNG still experiences slow uptake due to the lack of bunkering infrastructure. Identifying economic incentives associated with the provision of such infrastructure, in the form of a competitive advantage for ports, would contribute to better adoption of this fuel, and could be accompanied by policy support.

### 2.2. Can ports expect a gain in attractiveness from LNG infrastructure?

The attractiveness of a port reflects its ability to be selected by carriers for shipments (Malchow and Kanafani, 2004). Several

<sup>1</sup> The marginal social benefits of BEV represent “the incremental climate benefits due to the reduction of greenhouse gases. [...] [The social benefits of BEV] are proportional to the number of BEV adopters” (Brozynski and Leibowicz, 2022; pp 763-764).

<sup>2</sup> Regardless of the size of the company and the adoption status (early adopters, early followers, late followers, and laggards).

studies have evaluated the criteria that affect port choice decisions. While port authorities have the possibility to act on some of them, other factors are exogenous to their management strategy. In other words, ports cannot directly act on all factors of attractiveness (Lirn et al., 2004; Sánchez et al., 2011). Among the factors they cannot influence, geographical location, connectivity to cargo-generating hinterlands, and political stability are the most significant. Using disaggregate panel data for eight ports, Malchow and Kanafani (2004) found that port location variables, such as oceanic and inland distances from the origin of the shipment to the port, significantly affect port choice. A questionnaire of Lam et al. (2011) to ship operators showed the stability of the political environment to be the most important factor of competitiveness in the ports of Shanghai and Singapore.

In terms of the factors that ports can control, the most prominent are related to the quality of the infrastructure. First, ports can improve their attractiveness by providing adequate infrastructure to match the demands of carriers. The adequacy of such infrastructure is either based on the specificity of the vessels (bunkering stations, cargo types) or on their cargo handling capacity (Tongzon, 2009). Second, ports can improve the efficiency of their existing infrastructure. Port efficiency refers to the speed and reliability of their services (Sánchez et al., 2003) and is perceived as the key element of the quality of services and port attractiveness by stakeholders (Sánchez et al., 2011; Tongzon, 2009).

Other factors are the costs related to transportation, which include port fees and the operational travel costs (such as fuel consumption), and the port's reputation and reliability (Lam et al., 2011; Sánchez et al., 2003; Ugboma et al., 2006). Although travel costs can be a barrier to attractiveness in some shipping segments (such as short-sea and container shipping) (Nir et al., 2003), they are not the primary factor in most (D'Este and Meyrick, 1992; Murphy and Hall, 1995; Veldman et al., 2011). Moreover, they are difficult to incorporate in empirical analyses (Malchow and Kanafani, 2004). A port's reputation and reliability are derived from the perception of the overall quality of services (Sánchez et al., 2003).

Very few studies have considered the case of bunkering infrastructure for alternative fuels as a factor of port attractiveness. From a theoretical perspective, businesses, including ports, perceive environmental investments as a way to differentiate their operations toward being more green and to gain a competitive advantage (Aghion et al., 2023). Although bunkering infrastructure is a component of ports' corporate social responsibility (CSR) strategy, the effect of such a sustainability plan on competitiveness is not always evident. This can be explained by the absence of proper CSR metrics (Stein and Acciaro, 2020) and the inexistence of clear business case studies directly associating competitiveness of ports with their environmental performance (Adams et al., 2009). Stein and Acciaro (2020) showed that even when ports' corporate sustainability is well measured, its effect is only indirect through branding, reputation, safety, security, and risk management.

Among empirical studies, Valionienė and Strakauskaite (2015) assessed the effect of the establishment of an LNG terminal at the Klaipėda seaport (Lithuania). Based on a survey of shipping operators and bunkering providers, the authors found that the infrastructure could lead to increased attractiveness if the fuel is available in sufficient quantity and the bunkering and overall operations provided by the port are highly efficient.<sup>3</sup> Using a fuzzy analytical hierarchy process (fuzzy-AHP), Kim et al. (2021) concluded that the safety and security of the LNG-bunkering port were the most significant factors of attractiveness for the interviewed shipping companies. More generally, the results of these studies are in accordance with those of Lam et al. (2011), who identified, for the case of Singapore and Shanghai, a set of attributes that play a role in the competitiveness of bunkering ports (ranging from the quality, adequacy, and price of the bunkering services to the country's political transparency and stability). However, the authors insisted that the relative importance of these attributes varies from case to case. Furthermore, it is difficult to generalize the implications of these case studies as they were conducted on a small sample of ports.

Unlike existing studies based on surveys, we analyze the effective vessel movements in the prediction of port choice for the 135 Baltic ports that received calls from LNG-fueled vessels in 2019. We constructed a multilevel database from which the characteristics of both vessels and ports were included in regression modeling of port-vessel relationships in the form of choice probabilities. This helped us control for other factors that affect port choice while estimating the effect of LNG bunkering facilities.

Although most studies have modeled port choice using (multinomial) logit models (Anderson et al., 2009; Malchow and Kanafani, 2004; Veldman et al., 2011), we modeled the empirical probabilities ( $p_{ij}$ ) rather than discrete choices. The probability  $p_{ij}$  has the advantage of aggregating all port choices made during the timeframe and can also be interpreted as the loyalty of vessels to a set of ports (i.e., how often a given port was chosen during the studied timeframe). From the literature on port customer retention, loyalty is perceived as an indicator of port attractiveness (Jiannan et al., 2020; Yuen et al., 2018). Given the range of the values of  $p_{ij}$ , multilevel beta regressions were adopted, as they are suited for regression settings in which the outcome variables are proportions. Because of their characteristics, they have started gaining popularity in empirical studies in the shipping industry. For example, Chen et al. (2017) used beta regressions to model the effect of service attributes on customer retention in container shipping, while Acocella et al. (2020) estimated the effect of shipper-carrier relationships on the primary carrier acceptance ratio.<sup>4</sup>

### 3. Data and methodology

#### 3.1. Data

The explained variable is the probability ( $p_{ij}$ ) that vessel  $j$  chooses to call at port  $i$ . It is analogous to the market share (or share of

<sup>3</sup> Other factors include competitive price, geographical location, supplier reliability, and political stability. However, the survey respondents considered them less influential.

<sup>4</sup> Measured as the fraction of loads that are accepted by the primary carrier relative to the total number of loads offered from a specific shipper.

traffic) that port  $i$  is able to attract from vessel  $j$ , and is determined as the ratio

$$P_{ij} = \frac{Y_{ij}}{\sum_i Y_{ij}} \quad (1)$$

where  $Y_{ij}$  represents the number of calls made by vessel  $j$  at port  $i$ , or in other words, the number of times vessel  $j$  chose port  $i$  in the studied timeframe (year 2019).

We used vessel movement data to build the port choice model. The data were obtained from S&P's IHS Markit product, which lists all vessels that called in the BSR in 2019<sup>5</sup> and their characteristics, as well as the facilities of the ports of call. Our model relied on the assumption that vessels make their port selection based, not only on the ports' infrastructure and factors of attractiveness, but also on their own characteristics, such as the type of cargo. The vessel data included a classification based on the type of cargo (container ships, dry bulk carriers, passenger ships, RoRo, tankers, or general cargo ships), as well as the fuel used (e.g., LNG, distillate fuels, etc., or dual-fuel systems) which allowed us to select only vessels that use, at least in part, LNG for their engine. In terms of the port facilities, the data indicated whether the ports had adequate facilities (dummy variables) to receive passengers or to handle specific shipments, such as dry bulk, containers, vehicles (Ro-Ro facilities), and liquids (oil, chemicals, etc.). The movement data also provided the total time spent in port for each call. Based on this information, we were able to derive the vessel turnaround time, which is a good indicator of port efficiency (Tongzon, 2009). More precisely, as ports received multiple calls, sometimes from different vessels, the median turnaround time was considered to be a representative proxy of port efficiency.

Although the IHS Markit database references the Baltic ports with LNG facilities, no information is available concerning the facilities that are specifically involved in the bunkering of marine vessels. Indeed, several types of facilities, such as liquefaction plants,<sup>6</sup> are not directly part of the bunkering process. Thus, we used additional sources to identify the ports that offered specific bunkering services to vessels in 2019. The main additional source of data was the small-scale LNG database from Gas Infrastructure Europe (GIE).<sup>7</sup> For each port, the database references the type of LNG installations (bunkering ships, liquefaction plants, fuel loading ships or roads) and, when available, the name of the operator, the services offered, and the bunkering modes (ship-to-ship bunkering, truck-to-ship bunkering, or terminal tank to ship bunkering via a pipeline), as well as the start-up year. We exclusively considered ports with bunkering operations in 2019 and earlier. The other sources consisted of documents from ports, terminal operators or press releases for the ports for which the data from IHS Markit and GIE were either missing or unclear. Table 1 is a synthesis of the data description, while Table 2 gives details on the ports with LNG bunkering facilities. Summary statistics of the variables are available in the Appendix.

### 3.2. The multilevel nature of our data

Our model predicts the port choice probability for port  $i$  by vessel  $j$  depending on the characteristics of  $i$  and  $j$ . The outcome variable is observed at the port-vessel interaction level ( $ij$ ). That is, in such a structure, the choice probabilities are correlated between (i.e., clustered within) ports and vessels. It is said that such data have a multilevel structure, in which the port-vessel interaction ( $ij$ ) is the lowest level, and the port and vessel levels are the highest. In this situation, in addition to the variables included in the model, unobserved port-specific and vessel-specific effects may affect the probability of port choice. Such effects can be, for example, the ports' reputation or the carriers' preferences (Martínez-Pardo et al., 2020), respectively. To control such unobserved effects, the exact nature of the multilevel structure must be identified. Given the design of our study, two structures were possible, i.e., a nested or cross-classified structure.

Under a nested structure, e.g., ports nested within vessels, each port would receive calls from one unique vessel. Although this was the case for some ports in our sample (30.37 % of the ports, i.e., 41 of 135 ports<sup>8</sup>), this did not represent the general structure of our design (on average, each port had a relationship with 4.52 vessels). In addition, we could not assume a case in which vessels were nested within ports because in our sample, vessels operated at least between two ports (origin and destination). Under a cross-classified structure, there should be no hierarchical relationship between these levels (ports are not nested within vessels and vessels are not nested within ports). See Fig. 1 for an illustration, in which the level-1 units are the ports, the level-2 units are the vessels, and the lowest level of observations is from the vessel port choice probabilities (the vessel-port level,  $ij$ ).

In our case, a cross-classified structure was the best design to model port-vessel relationships for two major reasons. The first benefit of a cross-classified structure is that it encompasses both nested and non-nested relationships, such as situations in which ports receive calls from more than one unique vessel and irregular port choice behaviors (e.g., container shipping and other footloose segments). Second, in the estimation of the effect of LNG bunkering infrastructure, which is measured at the port level, assuming only a nested structure while cross-classified relationships are present in the data can lead to incorrect inferences (Rasbash and Browne, 2008). The standard errors of the coefficients would be underestimated because not all sources of variation would have been included in the model. In some cases, such under-specification can wrongfully lead to the parameter of interest being considered significant. In

<sup>5</sup> Although the year 2019 does not give the latest mapping of existing port facilities, it represents the most recent pre-pandemic observations.

<sup>6</sup> Liquefaction plants are used to turn natural gas into its liquid form (LNG) for more efficient transportation. They represent a key infrastructure in LNG export terminals (e.g., the Kollsnes I and II terminals in Norway).

<sup>7</sup> GIE is an independent, non-profit European organization representing operators working in the gas industry with low-carbon and renewable molecules. Databases are open source and available at <https://www.gie.eu/transparency/databases/> (last consulted by the authors September 1, 2023).

<sup>8</sup> Note that our sample only includes LNG-fueled vessels, which explains why 41 ports accommodate only one vessel.

**Table 1**  
Data description.

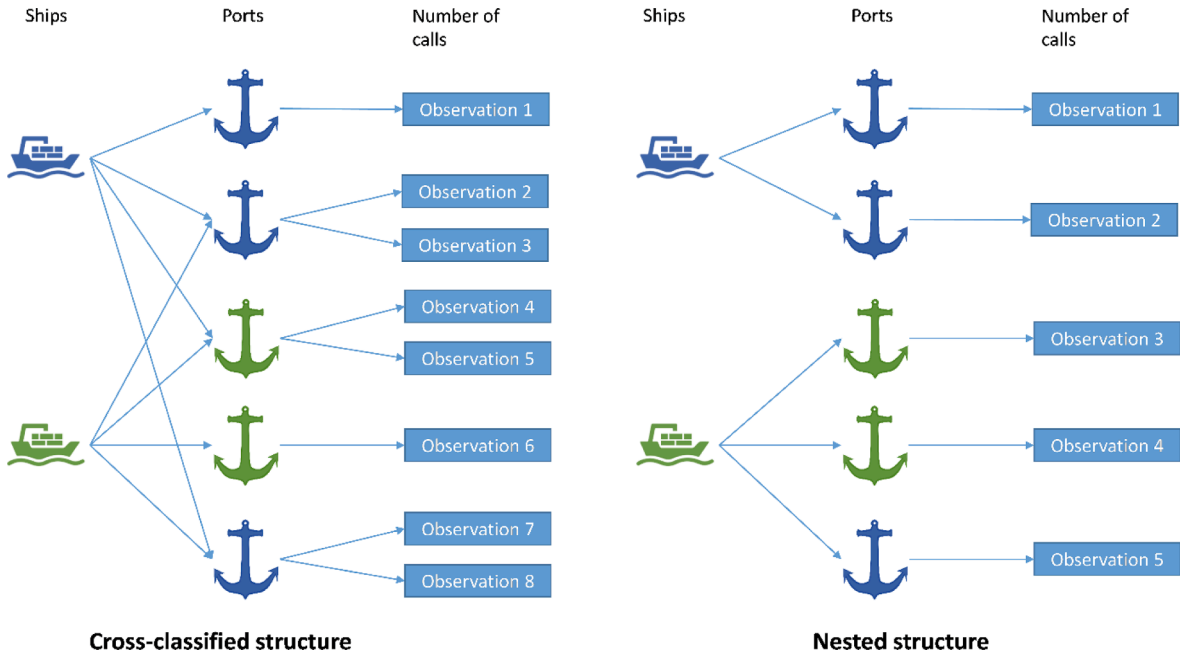
Variable name	Description	Source
Port choice probability $P_{ij}$	Rate of calls vessel $j$ has made in port $i$ in year 2019.	IHS Markit data
Ship type	The type of vessel (container ship, tanker ship, passenger ship, RoRo ship, General cargo ship, other bulk dry ship)	
Vessel turnaround, per ship type	Median time (in hours) spent by vessels in port. Measured at port level, per ship type	
Port country	Country of the port, in the Baltic	
Port facilities:		
Dry bulk facilities	Facilities available in port, 1 if yes; 0 if not	
Breakbulk facilities	Facilities available in port, 1 if yes; 0 if not	
Container facilities	Facilities available in port, 1 if yes; 0 if not	
RoRo facilities	Facilities available in port, 1 if yes; 0 if not	
Passenger facilities	Facilities available in port, 1 if yes; 0 if not	
Liquid facilities	Facilities available in port, 1 if yes; 0 if not	
LNG bunkering facilities	Facilities available in port, 1 if yes; 0 if not	GIE small-scale LNG database; terminal operators' website; <a href="#">LNG in Baltic Sea Ports (2014)</a> ; press releases.
LNG bunkering modes	Ship-to-ship (STS) and/or truck-to-ship (TTS) and/or tank to pipeline-to-ship (PTS)	

**Table 2**  
Ports in the Baltic Sea Region (BSR) with LNG bunkering facilities in 2019.

Port name	Country	Bunkering mode(s) available in 2019	Start-up year
Brofjorden	Sweden	Ship-to-ship (STS); tank to pipeline-to-ship (PTS) <u>Source:</u> <a href="#">European Commission (2014b)</a> ; <a href="#">LNG in Baltic Sea Ports (2014)</a>	2014
Gothenburg	Sweden	Ship-to-ship (STS); truck-to-ship (TTS) <u>Source:</u> <a href="#">Port of Gothenburg website, n.d.</a>	2015
Helsinki	Finland	Ship-to-ship (STS); truck-to-ship (TTS) <u>Source:</u> <a href="#">Port of Helsinki (2017)</a>	2014
Klaipeda	Lithuania	Tank to pipeline-to-ship (PTS) <u>Source:</u> <a href="#">AB Klaipėdos nafta website (n.d.)</a> ; <a href="#">European Commission (2014a)</a>	2015
Lysekil	Sweden	Ship-to-ship (STS) <u>Source:</u> GIE small-scale LNG database	2017
Nynäshamn	Sweden	Truck-to-ship (TTS); tank to pipeline-to-ship (PTS) <u>Source:</u> <a href="#">LNG in Baltic Sea Ports (2014)</a>	2011
Oskarshamn	Sweden	Ship-to-ship (STS); truck-to-ship (TTS) <u>Source:</u> <a href="#">HELCOM (2019)</a> ; <a href="#">Smålandshamn AB (2019)</a>	2019
Pori	Finland	Tank to pipeline-to-ship (PTS) <u>Source:</u> <a href="#">Gasum LNG (2021)</a>	2016
Risavika (Stavanger)	Norway	Ship-to-ship (STS); tank to pipeline-to-ship (PTS) <u>Source:</u> <a href="#">Offshore Energy (2015)</a> ; GIE small-scale LNG database	2015
Rostock	Germany	Ship-to-ship (STS); truck-to-ship (TTS) <u>Source:</u> GIE small-scale LNG database; <a href="#">Serry (2017)</a>	2016
Stockholm	Sweden	Ship-to-ship (STS); truck-to-ship (TTS) <u>Source:</u> <a href="#">LNG in Baltic Sea Ports (2014)</a> ; GIE small-scale LNG database	2013
Vysotsk	Russia	Tank to pipeline-to-ship (PTS) <u>Source:</u> <a href="#">Mampaey (2023)</a>	2019

Note: The sample only includes the ports that accommodated at least one LNG-powered vessel in 2019.

addition, the clustered variance (e.g., the random vessel-specific effects on port choice) from the nested structure could not be trusted due to its instability following the incorporation of additional factors in the model. The reader may refer to the studies of [De Leeuw et al. \(2008\)](#) and [Schielzeth and Nakagawa \(2013\)](#) for more distinctions between nested and cross-classified models and to that of [Rasbash and Browne \(2008\)](#) for the consequences of ignoring non-hierarchical structures. However, for the purpose of comparison, we also estimated a nested model in addition to the main cross-classified model. Our sample consisted of 73 vessels and 135 ports, resulting in 610 observations at the port-vessel ( $ij$ ) level.



**Fig. 1.** Schematic illustration of crossed-classified and nested structures in multilevel models. Source: Authors conception, adapted from Schielzeth and Nakagawa (2013). The green color represents LNG-fueled ships and ports with LNG bunkering infrastructure. The blue color represents oil-fueled ships and ports without LNG bunkering infrastructure. In the nested structure, ships can call only at specific ports. In the cross-classified structure, their port choice probabilities are a priori not null for the other ports, with or without LNG bunkering infrastructure.

### 3.3. Methodology

#### 3.3.1. Multilevel beta regressions

The beta distribution is a versatile family distribution for proportion data (taking values between 0 and 1) defined by two shape parameters. Ferrari and Cribari-Neto (2004) suggested the following parametrization over the mean ( $\mu$ ) and the precision ( $\phi$ ):

$$f(p; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} p^{\mu\phi-1} (1-p)^{(1-\mu)\phi-1} \quad (2)$$

where  $E(p) = \mu$ ,  $\text{Var}(p) = \frac{\mu(1-\mu)}{1+\phi}$ ,  $0 < p < 1$ , and  $\Gamma(\cdot)$  is the Gamma function.

In the regression sample, the occurrences  $p_{ij} = 0$  and  $p_{ij} = 1$  were discarded because  $p_{ij}$  must be continuously distributed in the interval  $]0, 1[$  (Ferrari and Cribari-Neto, 2004).

When dealing with proportion data, beta regressions present several advantages over linear regressions. Ordinary least squares (OLS) regressions lead to inconsistent predictions when the dependent variable is bounded, as they can, for example, predict negative proportions (which are inappropriate in our case). Furthermore, skewness and non-constant variances are well accommodated in the model, which makes it robust to heteroscedasticity (Ferrari and Cribari-Neto, 2004).

Noting  $LNG_i$ , the binary variable for the availability of LNG bunkering facilities in port  $i$  and  $\delta$  its coefficient (log-odds ratio), Eq. (3) represents the regression model

$$\mu_{ij} = \frac{\exp(X_{ij}\beta + \delta LNG_i + Z_{ij}b)}{1 + \exp(X_{ij}\beta + \delta LNG_i + Z_{ij}b)} \quad (3)$$

where  $\mu_{ij}$  denotes the mean probability under a logit link function.<sup>9</sup>  $X_{ij} = (1, x_{ij}^1, \dots, x_{ij}^K)$  is the (row) vector containing the covariates (explanatory variables) at both port ( $i$ ) and vessel ( $j$ ) levels;  $\beta = (\beta^0, \beta^1, \dots, \beta^K)^T$  is the vector of coefficients. The vector  $b \sim N(0, \sigma_b)$  contains the random effects and  $Z_{ij}$  is the matrix associated with each term, i.e. the port ( $i$ ) or vessel ( $j$ ) random component. More precisely, under the cross-classified model,  $Z_{ij} = z_i + z_j$  while  $Z_{ij} = z_j$  under the nested model.

Although the cross-classified multilevel regression accounts for confounders at both the vessel and port levels, some bias may

<sup>9</sup> A link function connects the mean of the explained variable ( $\mu_{ij}$  in the present case) to the explanatory variables in generalized linear models. Considering a logit link, equation (3) is derived from  $g(\mu_{ij}) = \log(\mu_{ij}/(1-\mu_{ij})) = X_{ij}\beta + \delta LNG_i + Z_{ij}b$

remain due to the endogeneity of the predictor variable of interest, i.e., LNG bunkering facilities in port. The availability of LNG bunkering facilities is likely to be endogenous as it is not randomly distributed between ports. In other words, ports decide (or not) to invest in such facilities. Following the literature on the evaluation of treatment effects, one would say that ports “self-select into treatment”. Such self-selection depends on a number of port characteristics, such as the country (not all countries in the Baltic have ports offering LNG bunkering to vessels), the types of vessels calling (e.g., tanker vessels), etc. (cf. Table 5 for differences between port characteristics). In cases in which these characteristics positively (negatively) correlate with the port choice probability, not controlling for self-selection would overestimate (underestimate) the effect of LNG bunkering. To circumvent such bias, we implemented multilevel propensity score matching (PSM) for a more robust estimation of the causal impact of LNG bunkering (Arpino and Mealli, 2011; Kim and Seltzer, 2007).

### 3.3.2. Multilevel propensity score matching (PSM)

The estimation of causal effects through matching methods is at the forefront in the literature on the evaluation of treatment effects.

The  $LNG_i$  variable introduced in Eq. (3) is the treatment in the present case. Note that  $p_{ij}^1$  and  $p_{ij}^0$  are the outcome (port choice probability) for the ports with and without LNG bunkering, respectively. Introducing counterfactual<sup>10</sup> notation, the observed difference in the average port choice is written as:

$$\begin{aligned} E(p_{ij}^1) - E(p_{ij}^0) &= E(p_{ij}^1 | LNG_i = 1) - E(p_{ij}^0 | LNG_i = 0) \\ &= \underbrace{E(p_{ij}^1 | LNG_i = 1) - E(p_{ij}^0 | LNG_i = 1)}_{\text{Causal effect}} + \underbrace{E(p_{ij}^0 | LNG_i = 1) - E(p_{ij}^0 | LNG_i = 0)}_{\text{Selection bias}} \end{aligned} \quad (4)$$

The first part of Eq. (4) represents, for the ports with LNG bunkering facilities, the causal effect of these facilities on their choice probability (the average treatment effect on the treated or ATT). It is the average difference between their actual outcome ( $E(p_{ij}^1 | LNG_i = 1)$ ) and what they would have had without LNG bunkering facilities, also known as the counterfactual. The second term measures the selection bias, which is the reason why the difference in the observed outcomes may not directly estimate the causal effect of LNG bunkering.

In the present context of non-randomized treatment assignment (Rubin, 1974), econometricians rely on two major assumptions to cancel out the bias and retrieve the ATT, these are the so-called conditional independence assumption (CIA) and the common support assumption. Under the CIA, also known as the *unconfoundedness* assumption, the outcome is independent of the treatment status when conditioning on the characteristics  $X_{ij}$ , or on a balancing score of  $X_{ij}$  such as the propensity score  $e(X_{ij})$  (Rosenbaum and Rubin, 1983), making the treatment assignment equivalent to a randomized setting (Eq. (5))

$$\{p_{ij}^1, p_{ij}^0\} \perp LNG_i | e(X_{ij}) \quad (5)$$

The propensity score,  $e(X_{ij}) = \Pr(LNG_i = 1 | X_{ij})$ , is defined as the probability for a unit to receive the treatment, given its set of observed covariates.

The common support (or overlap) assumption rules out the possibility of perfect predictability and imposes that individuals with the same covariates have the same probability  $0 < e(X_{ij}) < 1$  of being both treated and non-treated (Caliendo and Kopeinig, 2008; Heckman et al., 1999).

Our propensity scores were obtained from the multilevel logit model (Gelman and Hill, 2006; Goldstein, 2011; Kim and Seltzer, 2007) to accommodate the structure of our data. Arpino and Mealli (2011) showed that the two-level logit models with random intercepts (as implemented in the present article) perform better, in terms of bias correction and mean square error, than models omitting the multilevel nature of the data, and are suitable for estimation of the ATT.

As  $e(X_{ij})$  is a continuous variable, it is difficult for two units to share the exact same value for the propensity score. Matching is thus implemented using the nearest neighbor algorithm, with a caliper<sup>11</sup> of 0.01 around the propensity score (Kim and Seltzer, 2007). The matching is performed with replacement, meaning that control units can be matched with more than one treatment unit.

When all assumptions are met, the observed factors determining port choice are held constant and the only difference in the outcome stems from the availability or not of LNG bunkering facilities. Formally, the selection bias becomes null as  $E(p_{ij}^0 | e(X_{ij}), LNG_i = 1) = E(p_{ij}^0 | e(X_{ij}), LNG_i = 0)$ , and the causal effect (the ATT) is obtained as  $ATT = E(p_{ij}^1 - p_{ij}^0 | e(X_{ij}), LNG_i = 1)$ , i.e. the average difference in the port choice probabilities between the matched treatment and control groups. This difference can be assessed via

<sup>10</sup> The counterfactual represents the outcome that ports without LNG bunkering facilities would have had if they had LNG bunkering facilities, i.e.  $p_{ij}^0 | LNG_i = 1$ . Note that while  $p_{ij}^0 = (p_{ij}^0 | LNG_i = 0)$  and  $p_{ij}^1 = (p_{ij}^1 | LNG_i = 1)$  are observable values, the counterfactual cannot be observed with the data but can be retrieved with the PSM.

<sup>11</sup> The caliper represents the range of propensity scores  $e(X_{ij})$  within which the units will be matched (their  $e(X_{ij})$  are deemed sufficiently close). We set the caliper to 0.01, i.e., the range was  $e(X_{ij}) \pm 0.01 \times \text{std.dev.}(e(X_{ij}))$ .

comparison of the means (e.g., *t*-test) or regression. Angrist and Pischke (2008) argued that the difference between the two approaches, if any, is minor.<sup>12</sup> However, in the presence of multilevel data, although adjustments such as the paired *t*-test or repeated measures ANOVA can control for port choice correlations within vessels or ports, multilevel regression models perform better (Ashbeck and Bell, 2016; Huber et al., 2015; Yu et al., 2022). Thus, the ATT is estimated through multilevel beta regression on a sample composed of the treatment and control groups. The analysis was conducted using the R software packages *mgcv* (Wood, 2017) for the beta regressions and *MatchIt* (Ho et al., 2011) for the propensity score matching.

## 4. Results

### 4.1. The benchmark multilevel beta regressions

The estimates from the multilevel beta regressions are presented in Table 3, in which the first and second group of columns assume a cross-classified and nested data structure, respectively. The port choice probability was predicted on the full sample, i.e. without matching. We used a logit link function in the beta regression. Thus, the coefficients estimated in the table represent the effect of a one-unit change in the variables on the log-odds ratio of the port choice probability  $p_{ij}$  (not on the probability  $p_{ij}$  itself).

The results of the nested model are consistent with those from existing research on port choice, which identified infrastructure adequacy and port efficiency as significant factors that influence port choice. In particular, passenger vessels tend to be the most loyal shipping segment when the ports provide adequate facilities. Passenger ships had 3.6-fold greater chances<sup>13</sup> of choosing a port with passenger facilities versus another vessel *vis-à-vis* a port with different facilities. In addition, they were deterred by port inefficiency. One additional hour in turnaround time in port was shown to decrease the port choice probability expressed by passenger vessels by 0.62%.<sup>14</sup> For container ports (i.e., with container handling facilities), (in)efficiency of operations also affected competitiveness, as shown by a decrease in their port choice probability by 1.78% for every additional hour. Furthermore, ports with container facilities tend to attract a higher number and a broader range of vessels than ports without such facilities. This is identified in the estimation as the main effect of the “Container facilities” variable was significant (log-odds ratio of 1.0262, odds ratio of 2.79) but not the interaction term with container ships. The availability of LNG bunkering facilities in port also had a positive effect, but it was only significant at the 10% level. At the 10% significance level, ports with LNG bunkering facilities had a 26.79% higher port choice probability than other ports, all else being equal. The other significant factors included the liquid and breakbulk facilities and being a tanker ship, as well as the port country fixed effects and vessel random effects. We did not summarize the fitted random effects in the regression output due to their high number (one random effect,  $\beta_j^0$ , was estimated for each vessel). In the present case, their contribution to the model was assessed following the significance of the associated *effective degrees of freedom* (edf), because the *mgcv* R package estimates the random effects as non-linear smooth terms using spline bases (Hastie and Tibshirani, 1987). Significant edf indicates a strong contribution of random effects, as it positively correlates with their variance (Wood, 2017).

In cases in which port choice was affected by other port-specific characteristics unobserved in our model (e.g., reputation, governance mode, pricing strategy, etc.), the nested estimations are not sufficient for inference, as they omit these variations and underestimate the standard-errors (Rasbash and Browne, 2008). Under the cross-classified assumption, the port-specific characteristics (random effects) were fitted and had a significant effect on port choice probability (cf. significant edf in the left column of Table 3). The effect of LNG bunkering facilities in port was no longer significant when port-specific unobserved features were controlled for. This indicates the existence of confounding effects that affect both the provision of LNG bunkering and port choice probabilities. This is also true, for example, for the effect of turnaround time on the choice probability of container ports. Vessels tended to value other features from container ports more than turnaround time. For passenger ships, the effect of turnaround time, as well as the adequacy of passenger facilities, was sufficiently robust to resist the incorporation of port random effects.

Although the effect of LNG bunkering was not significant in the cross-classified model, we acknowledge that the various types of bunkering facilities (or bunkering modes) differ in efficiency. We investigated possible heterogeneity in their effect on port attractiveness to LNG-fueled vessels. The results of the regressions that incorporate the different bunkering modes, i.e., ship-to-ship (STS), truck-to-ship (TTS), and tank to pipeline-to-ship (PTS) are presented in Table 4. Note that ports providing LNG bunkering can jointly offer this service under multiple modes. Hence, these variables were introduced into the model in addition to the main LNG bunkering binary variable without an omitted category. The effect of most variables remained similar to the previous regressions, except for the provision of LNG bunkering under the nested structure. In this case, the main effect of LNG bunkering was negative, but only significant at the 10% level. However, the ports providing LNG bunkering through a TTS mode were the most attractive, with a port choice probability 3.3 times<sup>15</sup> higher than other ports with similar characteristics. This was followed by the PTS mode, with an odds ratio of 3.13. STS bunkering did not have a significant effect on port choice. The effect of the bunkering modes was not sufficiently robust to withstand the incorporation of port random effects in the cross-classified model.

<sup>12</sup> See Angrist and Pischke (2008) chapter 3 *Making regression make sense*, section 3.3.1, *Regression meets matching*.

<sup>13</sup> The estimated coefficient (log-odds ratio) for the interaction between passenger ship and passenger facilities in port is 1.2777. The odds ratio is  $\exp(1.2777) = 3.5884$

<sup>14</sup> The estimated log-odds ratio for the interaction term Passenger facilities X Passenger ship turnaround X Passenger ship is  $-0.0062$ , the odds ratio is  $\exp(-0.0062)$  and the percentage change is  $(\exp(-0.0062) - 1) \times 100 = -0.6181\%$ . Note that the change is measured in % and not in percentage points.

<sup>15</sup> The estimated log-odds ratio is 1.1913, the odds ratio is  $\exp(1.1913) = 3.2914$ .

**Table 3**

Beta regression of the port choice probability under cross-classified and nested data structures.

A. parametric coefficients	Cross-classified		Nested	
	Estimate	Std. Error	Estimate	Std. Error
Intercept	-1.9553***	0.6320	-2.7511***	0.4795
LNG bunkering in port	-0.0345	0.2684	0.2374*	0.1343
Container ship	-0.4224	1.3018	1.3033	1.2009
Tanker ship	0.8851	0.6885	1.0297*	0.5744
Passenger ship	-0.1384	0.7635	0.5511	0.5849
RoRo ship	-0.9505	0.7872	-0.1491	0.6602
General cargo ship	0.3049	0.5748	0.6217	0.4963
Estonia (Country FE)	-0.2791	0.5419	0.0020	0.3790
Finland (Country FE)	0.2706	0.3477	0.7192***	0.2191
Germany (Country FE)	0.3325	0.3432	0.4387**	0.2142
Latvia (Country FE)	0.0520	0.5595	0.4117	0.3172
Lithuania (Country FE)	0.1993	0.6088	0.4191	0.3074
Norway (Country FE)	-0.2159	0.3170	0.4966**	0.1935
Poland (Country FE)	-0.3467	0.5812	0.2069	0.3416
Russia (Country FE)	0.7322*	0.4210	1.0858***	0.2291
Sweden (Country FE)	0.0915	0.2931	0.2336	0.1759
RoRo facilities	-0.1874	0.3422	-0.1033	0.2315
RoRo facilities X RoRo ship	0.5967	0.4819	0.3106	0.4053
RoRo facilities X RoRo ship turnaround X RoRo ship	0.0491***	0.0171	0.0283	0.0183
RoRo facilities X RoRo ship turnaround	0.0069	0.0103	0.0045	0.0064
Liquid facilities	0.0786	0.3252	0.3654*	0.2105
Liquid facilities X Tanker ship	-0.3053	0.4748	-0.1457	0.3876
Liquid facilities X Tanker ship turnaround X Tanker ship	-0.0087	0.0076	-0.0077	0.0061
Liquid facilities X Tanker ship turnaround	0.0058	0.0071	0.0115**	0.0049
Break bulk facilities	-0.5865*	0.3299	-0.7454***	0.2085
Container facilities	0.8775***	0.3040	1.0262***	0.2066
Container facilities X Container ship	0.8371	1.1905	-0.6802	1.1363
Container facilities X Container ship turnaround X Container ship	0.0186	0.0341	0.0291	0.0339
Container facilities X Container ship turnaround	-0.0150	0.0107	-0.0180***	0.0067
Passenger facilities	-0.4191*	0.2324	-0.1949	0.1367
Passenger facilities X Passenger ship	2.1558***	0.5698	1.2777***	0.3939
Passenger facilities X Passenger ship turnaround X Passenger ship	-0.0066**	0.0029	-0.0062**	0.0030
Passenger facilities X Passenger ship turnaround	0.0012	0.0015	0.0007	0.0008
Dry bulk facilities	0.0194	0.2972	0.2521	0.1906
B. smooth terms	edf	Ref.df	edf	Ref.df
Port random effects (random intercept)	61.8185***	134.0000		
Vessel random effects (random intercept)	58.6938***	72.0000	54.5437***	67.0000

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

Although the cross-classified multilevel regressions accounted for confounders, at both the vessel and port levels, PSM was also used to control for the selection bias based on observed characteristics, and improved the causal interpretation of the effect of LNG bunkering on port choice probability.

#### 4.2. The effect of LNG bunkering on port choice for the matched sample

The output of the multilevel logit regression<sup>16</sup> predicting the probability of providing LNG bunkering facilities is presented in Table A1 of the Appendix.

The observations that were matched in the treated and control groups, based on their propensity score, are plotted in Fig. 2. The figure confirms the common support (overlap) assumption between the treated and control groups. Observational units from ports with and without LNG bunkering facilities now show the same conditional probability of providing LNG bunkering facilities.

For the CIA, the balance of the covariates between the treatment and control groups was compared by measuring the standardized mean difference (Table 5). Although PSM improved the standardized mean difference of the covariates, the matching was not perfect

<sup>16</sup> When predicting the probability of LNG bunkering in ports, only vessel-specific random effects are fitted. This is because the port level is the lowest observational level and random effects cannot be estimated for ports. Nevertheless, the regression estimating the ATT incorporates both random effects if matched units are nested within ports and vessels.

**Table 4**  
Beta regression of the port choice probability including LNG bunkering modes.

A. parametric coefficients	Cross-classified		Nested	
	Estimate	Std. Error	Estimate	Std. Error
Intercept	−2.0039***	0.6409	−2.7856***	0.4919
LNG bunkering in port	−1.4241	0.9855	−0.9364*	0.5189
Ship-to-ship (STS) bunkering	0.2331	0.5294	0.0946	0.2096
Truck-to-ship (TTS) bunkering	1.1445	0.7230	1.1913***	0.4455
Pipeline-to-ship (PTS) bunkering	1.3439	0.8484	1.1395**	0.4751
Container ship	−0.4054	1.3094	1.5277	1.2131
Tanker ship	1.0418	0.6954	1.3692**	0.5816
Passenger ship	−0.1616	0.7750	0.5285	0.5922
RoRo ship	−0.9921	0.7948	−0.2125	0.6736
General cargo ship	0.2908	0.5803	0.6122	0.5089
Estonia (Country FE)	−0.2916	0.5446	0.0102	0.3780
Finland (Country FE)	0.3000	0.3526	0.7420***	0.2228
Germany (Country FE)	0.3411	0.3472	0.4228**	0.2140
Latvia (Country FE)	0.0148	0.5654	0.3689	0.3164
Lithuania (Country FE)	0.2881	0.6437	0.4849	0.3250
Norway (Country FE)	−0.2280	0.3200	0.4956**	0.1938
Poland (Country FE)	−0.3437	0.5862	0.2015	0.3445
Russia (Country FE)	0.7510*	0.4287	1.1020***	0.2327
Sweden (Country FE)	0.0696	0.2979	0.1862	0.1789
RoRo facilities	−0.3612	0.3582	−0.2981	0.2420
RoRo facilities X RoRo ship	0.6377	0.4859	0.3967	0.4070
RoRo facilities X RoRo ship turnaround X RoRo ship	0.0491***	0.0171	0.0282	0.0183
RoRo facilities X RoRo ship turnaround	0.0075	0.0106	0.0064	0.0065
Liquid facilities	0.0448	0.3302	0.3279	0.2132
Liquid facilities X Tanker ship	−0.4602	0.4799	−0.4566	0.3827
Liquid facilities X Tanker ship turnaround X Tanker ship	−0.0089	0.0077	−0.0083	0.0062
Liquid facilities X Tanker ship turnaround	0.0051	0.0072	0.0117**	0.0049
Break bulk facilities	−0.4730	0.3396	−0.6734***	0.2092
Container facilities	0.9393***	0.3126	1.1440***	0.2105
Container facilities X Container ship	0.8092	1.1946	−0.9101	1.1402
Container facilities X Container ship turnaround X Container ship	0.0187	0.0340	0.0296	0.0339
Container facilities X Container ship turnaround	−0.0155	0.0110	−0.0195***	0.0068
Passenger facilities	−0.3573	0.2441	−0.1567	0.1453
Passenger facilities X Passenger ship	2.1925***	0.5812	1.3612***	0.3920
Passenger facilities X Passenger ship turnaround X Passenger ship	−0.0066**	0.0029	−0.0064**	0.0029
Passenger facilities X Passenger ship turnaround	0.0013	0.0015	0.0009	0.0008
Dry bulk facilities	0.0883	0.3042	0.2798	0.1934
B. smooth terms	edf	Ref.df	edf	Ref.df
Port random effects (random intercept)	60.3432***	134.0000		
Vessel random effects (random intercept)	58.8425***	72.0000	55.1980***	67.0000

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

and the imbalance remained. However, using simulations data, [Nguyen et al. \(2017\)](#) showed that the residual imbalance can be markedly reduced by including all covariates with a standardized mean difference  $> 0.10$  in the ATT regression. Thus, we consider the CIA to be met.

The regression outputs of the causal effect of LNG bunkering (ATT) with and without adjustment for unbalanced covariates are presented in [Tables 6 and 7](#), respectively. PSM matched observations within and across ports and vessels and thus, the ATT regression controlled for port and vessel random effects. The regression in [Table 6](#) controlled for the presence of RoRo ships, for the availability of liquid facilities and for country fixed effects, as a balance was not attained for these variables. Under this setting, the effect of LNG bunkering in port was not significant. Note that none of the included covariates had a significant effect, but the intercept and random effects were significant. This result highlights that although these covariates had a standardized mean difference above the 0.10 threshold ([Nguyen et al., 2017](#)), they were not at the origin of the remaining difference in port choice probabilities between ports with and without LNG bunkering facilities. Furthermore, when the regression did not control for these covariates ([Table 7](#)), LNG bunkering had a significant negative impact on port choice probability, which is biased as confounded with the residual imbalance. We also used PSM and ATT regression estimations when considering each of the bunkering modes as the treatment. The three alternative estimations led to the same conclusion as for the global LNG bunkering variable, i.e. no significant ATT was found. These regression outputs can be found in the [Appendix](#).

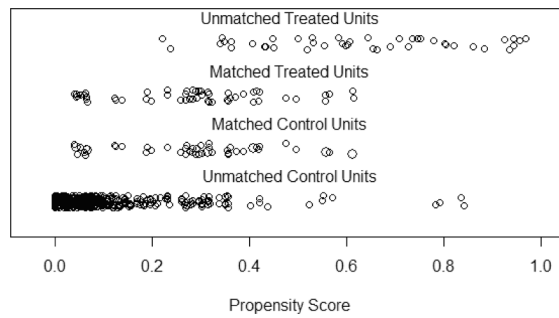


Fig. 2. Distribution of the propensity scores and assignment to the treatment and control groups.

**Table 5**  
Covariates balance before and after propensity score matching.

	Before propensity score matching			After propensity score matching		
	Treated	Control	Std.Mean Diff.	Treated	Control	Std.Mean Diff
Container ship	0.073	0.039	0.131	0.019	0.019	0
Tanker ship	0.688	0.564	0.266	0.736	0.755	-0.041
Passenger ship	0.073	0.111	-0.146	0.038	0.057	-0.073
RoRo ship	0.062	0.101	-0.160	0.038	0	0.156
General.Cargo ship	0.083	0.097	-0.050	0.132	0.151	-0.068
Denmark	0	0.111	-0.385	0	0	0
Estonia	0	0.016	-0.137	0	0	0
Finland	0.135	0.078	0.168	0.132	0.189	-0.165
Germany	0.052	0.086	-0.151	0.094	0.075	0.085
Latvia	0	0.027	-0.182	0	0	0
Lithuania	0.115	0.006	0.341	0.019	0	0.059
Norway	0.208	0.391	-0.450	0.226	0.132	0.232
Poland	0	0.054	-0.261	0	0	0
Russia	0.052	0.068	-0.072	0.075	0.075	0
Sweden	0.438	0.163	0.552	0.453	0.528	-0.152
RoRo Facilities	0.771	0.792	-0.050	0.830	0.868	-0.090
Liquid Facilities	0.969	0.862	0.614	0.943	0.962	-0.108
Breakbulk Facilities	0.823	0.926	-0.270	0.962	0.943	0.049
Container Facilities	0.708	0.739	-0.068	0.774	0.792	-0.042
Passenger Facilities	0.531	0.716	-0.370	0.604	0.604	0
DryBulk Facilities	0.823	0.887	-0.168	0.962	0.943	0.049

**Table 6**  
LNG bunkering ATT regression.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept	-1.9536*	0.9973	-1.9589	0.0584
LNG bunkering in port	-0.2697	0.4687	-0.5755	0.5688
RoRo ship	-0.1173	0.9639	-0.1217	0.9039
Germany (Country FE)	0.3042	0.8202	0.3708	0.7131
Lithuania (Country FE)	0.6139	1.4596	0.4206	0.6767
Norway (Country FE)	0.9597	0.8070	1.1893	0.2426
Russia (Country FE)	1.4025	0.8546	1.6412	0.1100
Sweden (Country FE)	0.4437	0.5887	0.7537	0.4562
Liquid facilities	-0.6754	0.8614	-0.7841	0.4384
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	25.7620***	38.0000	348.5218	0.0001
Vessel random effects (random intercept)	31.3676***	38.0000	230.3248	0.0016

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

**Table 7**  
LNG bunkering ATT regression without covariates.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
LNG bunkering in port	-1.6007***	0.5508	-2.9061	0.0065
B. smooth terms				
	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	33.2242*	38.0000	1251.4046	0.0750
Vessel random effects (random intercept)	33.2986	39.0000	4777.5220	0.1730

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

The initial covariate imbalance between ports with and without LNG bunkering showed that ports decide to provide this infrastructure based on their observed characteristics (and those of the vessels that call). The results of the PSM and multilevel beta regressions showed that, beyond the observed characteristics, the two categories of port differ intrinsically, and this difference affects their port choice probabilities. After controlling for all those factors, we cannot conclude that there is a significant effect of LNG bunkering facilities on port competitiveness.

#### 4.3. Discussions on the methodology

Despite the efficiency of PSM in reducing sample selection bias, it presents certain limitations (King and Nielsen, 2019). First, PSM has been criticized because it might fail to balance the sample covariates. The main advantage of PSM is that, instead of matching for each of the  $K$  covariates, it overcomes the dimensionality curse by reducing the multivariate variations to a scalar, the propensity score, on which the matching is based. However, in cases in which the number of covariates is high, narrowing the dimensions may fail to effectively balance the samples (Brookhart et al., 2006; King and Nielsen, 2019). This was only a minor issue in our methodology, as we examined the standardized mean difference of each covariate after matching (Table 5) and adjusted for the remaining imbalance in the ATT regression, as suggested by Nguyen et al. (2017). The second criticism of PSM is that it may lead to biased estimates in cases of misspecification in the propensity score equation (King and Nielsen, 2019; Smith and Todd, 2005). Thus, we implemented an alternative matching method, Mahalanobis distance matching (MDM). MDM matches the covariates based on the Mahalanobis distance, which is a standardized version of the Euclidian distance. MDM does not require prior regression specification. Its application to our dataset yielded a similar residual imbalance as PSM, and the ATT regression, conditional on the unbalanced covariates, resulted in a non-significant effect of LNG bunkering on port competitiveness, as for PSM. The results are available upon request.

In addition to the limitations of PSM, the non-exhaustive characterization of LNG bunkering facilities and the uncertainty associated with the prominence of dual-fuel vessels should be considered, as well as the fact that our port choice model was only a partial approach to supply chains. First, we did not consider the service quality of LNG bunkering infrastructure in our evaluation. However, service quality is likely to correlate with the bunkering modes that we considered (ship-to-ship, truck-to-ship, and terminal tank to ship via pipeline). Differences in the bunkering capacity, pumping rate, or price of the bunkering could reveal heterogeneous effects. Such data were not available for our study but further research should investigate for possible heterogeneous effects with respect to these specificities of bunkering.

Second, almost all vessels in our sample were equipped with a dual-fuel system (LNG for the main engine and a distillate fuel for the secondary). It is possible that these vessels were operating on distillate fuels more frequently than on LNG, resulting in a low bunkering rate for LNG and explaining the absence of attractiveness toward LNG-bunkering ports. We did not have access to data on the vessels' effective fuel consumption, thus cannot rule out this hypothesis. Future studies should investigate this possibility, especially as the prominence of dual-fuel systems raises questions about the adoption of alternative fuels over conventional bunker fuels.

Finally, our port choice model only considered the port-ship interface and did not account for the origin, or final destination of the cargo. Haralambides (2015) and Magala and Sammons (2008) advocated for port choice models that consider the full supply-chain and multi-modal transportation in the assessment of port attractiveness. However, in addition to the lack of data in our case, it would have been cumbersome to integrate these dimensions for all the port calls in the BSR for the year 2019.

## 5. Conclusions and implications

The adoption of alternative marine fuels could substantially contribute to the mitigation of shipping emissions. Little progress has been made toward this transition due to hesitation from ports in the provision of infrastructure, on the one hand, and from ship-owners in equipping their vessels, on the other. While most research on this so-called chicken-and-egg problem has highlighted the expected role of the ports, very few studies were supported by empirical evidence. Ports would endorse this role as part of a sustainability strategy to meet IMO regulations or the United Nations Sustainable Development Goals (Alamouh et al., 2022). However, there is little evidence of an economic incentive through a gain in attractiveness. Our results at the level of the BSR indicate that the ports' cargo handling infrastructure, efficiency, and unmeasured intrinsic features affect their competitiveness, but the LNG bunkering facilities do not. The implications of our findings are twofold. First, they reveal the difficulty to generalize the results of case-specific studies, as

they cannot isolate the effects of the ports' intrinsic features that were difficult to measure or not included in the model. Second, our findings corroborate, to some extent, the hesitation of ports, which are still lagging in the provision of alternative fuel infrastructure as such investments do not guarantee a gain in attractiveness as argued in the literature.

Our results may nonetheless be subject to evolution in the future. We considered the case of ports in the BSR in 2019, when LNG, despite being the most utilized alternative marine fuel, was still in the formative phase of its adoption process (i.e., 1 % of the fleet in the BSR was powered by LNG and 8 % of the Baltic ports provided bunkering services). Such a low level of LNG adoption may explain our results. [Leibowicz \(2018\)](#) highlighted for infrastructure-dependent technologies that the diffusion of infrastructure precedes the adoption of vehicles, which precedes the travel diffusion process. [Struben and Sterman \(2008\)](#) showed the existence of a critical threshold for the self-sustained adoption of alternative technologies. Hence, we may expect that, beyond a critical threshold of adoption, LNG infrastructure could affect port attractiveness and moderate the effects of the classical facilities ([Parola et al., 2017](#)). Furthermore, it cannot be excluded that the effect of LNG bunkering is simply a long-term process, rather than a short-term effect. Route and port choice are sometimes long-term decisions for carriers, the short-term decision being the assignment of shipments to vessels, which implicitly assigns shipments to ports ([Malchow and Kanafani, 2004](#)). In this case, carriers may take several years before adjusting port selection to the attractiveness of LNG-bunkering ports, granting them a green reputation and favorable market positioning in the longer term. Although LNG is perceived by most actors as a transitional fuel because of its fossil origin and methane slips issue ([Englert et al., 2021](#); [Wechsler, 2021](#)), similar analyses and methodologies could be applied to other alternative fuels, including hydrogen, methanol, or ammonia in the near future.

In this perspective, the development of alternative fuels bunkering infrastructure with relatively limited short-term benefits for ports but potential long-term societal gains calls for a combination of market and non-market strategies by ports ([Baron, 1995](#)). Ports could individually seek to be among the first to develop this type of infrastructure while collectively carrying out actions aimed at obtaining favorable public policies (subsidies, tax incentives, etc.) from national and international authorities (European in particular), thus promoting a form of regulatory co-creation ([Gao and McDonald, 2022](#)).

## Funding

This study was supported by the ShipTRASE project (Belmont Forum via ANR Grant ANR-20-BFOC-0003).

## CRedit authorship contribution statement

**Akoh Fabien Yao:** Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Maxime Sèbe:** Writing – review & editing, Validation, Data curation. **Laura Recuero Virto:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Abdelhak Nassiri:** Writing – review & editing, Methodology, Formal analysis. **Hervé Dumez:** Resources, Validation, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Summary statistics of the data

**Table A1**

Summary statistics of the port-level variables.

Statistic	N <sub>i</sub>	Mean	St. Dev.	Min	Max	q0.25	Median	q0.75
Average %calls ( $p_i$ )	135	0.107	0.117	0.001	0.583	0.026	0.063	0.154
LNG bunker	135	0.089	0.286	0	1	0	0	0
Ship to Ship	135	0.067	0.250	0	1	0	0	0
Truck to Ship	135	0.044	0.207	0	1	0	0	0
Pipeline to Ship	135	0.044	0.207	0	1	0	0	0
Denmark	135	0.119	0.324	0	1	0	0	0
Estonia	135	0.030	0.170	0	1	0	0	0
Finland	135	0.096	0.296	0	1	0	0	0
Germany	135	0.104	0.306	0	1	0	0	0
Latvia	135	0.022	0.148	0	1	0	0	0
Lithuania	135	0.015	0.121	0	1	0	0	0
Norway	135	0.319	0.468	0	1	0	0	0
Poland	135	0.030	0.170	0	1	0	0	0
Russia	135	0.044	0.207	0	1	0	0	0
Sweden	135	0.222	0.417	0	1	0	0	0
Dry bulk Facilities	135	0.859	0.349	0	1	1	1	1

(continued on next page)

**Table A1** (continued)

Statistic	N <sub>i</sub>	Mean	St. Dev.	Min	Max	q0.25	Median	q0.75
Breakbulk Facilities	135	0.874	0.333	0	1	1	1	1
Container Facilities	135	0.570	0.497	0	1	0	1	1
RoRo Facilities	135	0.689	0.465	0	1	0	1	1
Passenger Facilities	135	0.593	0.493	0	1	0	1	1
Liquid Facilities	135	0.748	0.436	0	1	0.5	1	1
Turnaround Tanker	135	20.970	20.596	0.500	173.000	9.876	16.500	25.250
Turnaround Passenger	135	32.340	78.269	0.000	771.000	10.139	20.189	31.638
Turnaround RoRo	135	20.556	16.882	1.000	98.039	8.003	18.248	28.588
Turnaround Container	135	22.326	17.174	0.500	98.039	10.875	19.108	28.951

**Table A2**

Summary statistics of the vessel-level variables.

Statistic	N <sub>j</sub>	Mean	St. Dev.	Min	Max	q0.25	Median	q0.75
Container ship	73	0.055	0.229	0	1	0	0	0
Tanker ship	73	0.493	0.503	0	1	0	0	1
Passenger ship	73	0.260	0.442	0	1	0	0	1
RoRo ship	73	0.068	0.254	0	1	0	0	0
General cargo ship	73	0.096	0.296	0	1	0	0	0
Other bulk dry ship	73	0.041	0.200	0	1	0	0	0

**Table A3**

Summary statistics of the variables, measured at the port-vessel interaction.

Statistic	N <sub>ij</sub>	Mean	St. Dev.	Min	Max	q0.25	Median	q0.75
%calls ( $p_{ij}$ )	610	0.120	0.164	0.0003	0.998	0.022	0.050	0.136
LNG bunker	610	0.157	0.364	0	1	0	0	0
Ship to Ship	610	0.149	0.357	0	1	0	0	0
Truck to Ship	610	0.061	0.239	0	1	0	0	0
Pipeline to Ship	610	0.097	0.296	0	1	0	0	0
Denmark	610	0.093	0.291	0	1	0	0	0
Estonia	610	0.013	0.114	0	1	0	0	0
Finland	610	0.087	0.282	0	1	0	0	0
Germany	610	0.080	0.272	0	1	0	0	0
Latvia	610	0.023	0.150	0	1	0	0	0
Lithuania	610	0.023	0.150	0	1	0	0	0
Norway	610	0.362	0.481	0	1	0	0	0
Poland	610	0.046	0.209	0	1	0	0	0
Russia	610	0.066	0.248	0	1	0	0	0
Sweden	610	0.207	0.405	0	1	0	0	0
Container ship	610	0.044	0.206	0	1	0	0	0
Tanker ship	610	0.584	0.493	0	1	0	1	1
Passenger ship	610	0.105	0.307	0	1	0	0	0
RoRo ship	610	0.095	0.294	0	1	0	0	0
General cargo ship	610	0.095	0.294	0	1	0	0	0
Other bulk dry ship	610	0.082	0.275	0	1	0	0	0
RoRo Facilities	610	0.789	0.409	0	1	1	1	1
Liquid Facilities	610	0.879	0.327	0	1	1	1	1
Breakbulk Facilities	610	0.910	0.287	0	1	1	1	1
Dry Bulk Facilities	610	0.877	0.329	0	1	1	1	1
Container Facilities	610	0.734	0.442	0	1	0	1	1
Passenger Facilities	610	0.687	0.464	0	1	0	1	1
Turnaround Tanker	610	17.704	16.936	0	173.000	8.000	16.000	24.000
Turnaround Passenger	610	29.238	93.175	0	771.000	0.000	2.000	24.505
Turnaround RoRo	610	13.884	14.634	0	91.484	2.000	10.000	22.541
Turnaround Container	610	13.636	14.439	0	91.484	0.000	10.878	21.000

**B. Regression output of the main multilevel logit (propensity score): LNG bunkering as treatment****Table B1**

Logit regression predicting the bunkering of LNG in ports (Propensity score).

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept	-20.1629	829.8838	-0.0243	0.9806
Container ship	2.4821***	0.9604	2.5846	0.0098

(continued on next page)

**Table B1** (continued)

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Tanker ship	1.3075*	0.6870	1.9031	0.0570
Passenger ship	1.5914**	0.7645	2.0816	0.0374
RoRo ship	1.3670*	0.8180	1.6710	0.0947
General cargo ship	1.4055*	0.7724	1.8197	0.0688
Estonia (Country FE)	-1.0930	2215.3788	-0.0005	0.9996
Finland (Country FE)	16.8122	829.8835	0.0203	0.9838
Germany (Country FE)	15.9689	829.8836	0.0192	0.9846
Latvia (Country FE)	-0.2376	1869.8681	-0.0001	0.9999
Lithuania (Country FE)	19.4390	829.8837	0.0234	0.9813
Norway (Country FE)	16.1157	829.8834	0.0194	0.9845
Poland (Country FE)	-0.1460	1448.3446	-0.0001	0.9999
Russia (Country FE)	15.8654	829.8836	0.0191	0.9847
Sweden (Country FE)	17.9414	829.8834	0.0216	0.9828
RoRo facilities	1.7195***	0.5514	3.1185	0.0018
Liquid facilities	1.2464*	0.6890	1.8091	0.0704
Break bulk facilities	-2.2837**	1.0228	-2.2329	0.0256
Container facilities	0.3376	0.4427	0.7627	0.4456
Passenger facilities	-1.6685***	0.3653	-4.5670	< 0.0001
Dry bulk facilities	0.8755	1.0051	0.8711	0.3837
B. smooth terms	edf	Ref.df	F-value	p-value
Vessel random effects (random intercept)	1.2043	67.0000	1.2408	0.4075

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

### C. Regression output after PSM with bunkering modes as treatment

#### Ship-to-ship bunkering as treatment

**Table C1**

Ship-to-ship bunkering ATT regression.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept	-2.2678*	1.1863	-1.9118	0.0611
Ship-to-ship LNG bunkering in port	-0.2044	0.5168	-0.3954	0.6941
Tanker ship	-0.6157*	0.3330	-1.8491	0.0698
General cargo ship	0.2352	0.3638	0.6463	0.5207
Germany (Country FE)	-0.5393	1.0076	-0.5352	0.5947
Norway (Country FE)	0.3297	0.7032	0.4689	0.6410
Sweden (Country FE)	-0.1792	0.7185	-0.2495	0.8039
Liquid facilities	0.2271	1.0565	0.2149	0.8306
Passenger facilities	0.4246	1.1633	0.3650	0.7165
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	13.4674	24.0000	13.2672	0.9581
Vessel random effects (random intercept)	20.4797	42.0000	3.1900	0.9966

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

**Table C2**

Ship-to-ship LNG bunkering ATT regression, without covariates.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Ship-to-ship LNG bunkering in port	-2.1986***	0.7098	-3.0972	0.0032
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	21.6187	24.0000	53.5432	0.2345
Vessel random effects (random intercept)	25.3506	42.0000	38.0065	0.9914

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the

random effects. The Ref.df term denotes the reference degrees of freedom.

#### Truck-to-ship bunkering as treatment

**Table C3**

Truck-to-ship bunkering ATT regression.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept	-1.6162	1.1054	-1.4620	0.1583
Truck-to-ship LNG bunkering in port	0.2728	0.5670	0.4811	0.6353
General cargo ship	-0.5617	0.7390	-0.7600	0.4555
Germany (Country FE)	-1.7378	1.3352	-1.3015	0.2069
Sweden (Country FE)	-1.2105	0.7401	-1.6355	0.1166
Container facilities	-0.3021	0.8528	-0.3543	0.7266
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	12.1512	20.0000	43.4966	0.9749
Vessel random effects (random intercept)	12.4418	18.0000	33.3441	0.9983

**Note:** The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

**Table C4**

Truck-to-ship LNG bunkering ATT regression, without covariates.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Truck-to-ship LNG bunkering in port	-2.5127**	1.0653	-2.3588	0.0288
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	18.0086	20.0000	177.9583	0.9535
Vessel random effects (random intercept)	13.2626	18.0000	264.0728	1.0000

**Note:** The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

#### Pipeline-to-ship bunkering as treatment

**Table C5**

Pipeline-to-ship bunkering ATT regression.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept	-1.0932	1.8506	-0.5907	0.6480
Pipeline-to-ship LNG bunkering in port	-0.6966	0.7534	-0.9246	0.5056
Tanker ship	1.2318	1.3458	0.9153	0.5090
RoRo ship	1.6018	1.4674	1.0916	0.4499
General cargo ship	-0.4898	1.4122	-0.3468	0.7805
Lithuania (Country FE)	-1.6426	1.4731	-1.1151	0.4429
Norway (Country FE)	-0.5298	1.1198	-0.4731	0.7089
Russia (Country FE)	-1.8138	1.2477	-1.4538	0.3568
Sweden (Country FE)	-1.7230	1.2310	-1.3998	0.3687
RoRo facilities	-4.9000	1.7885	-2.7397	0.1917
Container facilities	2.9345	1.4926	1.9661	0.2695
Passenger facilities	0.6731	1.2678	0.5309	0.6783
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	5.8778	13.0000	484895.3984	1.0000
Vessel random effects (random intercept)	6.9467	7.0000	58914.8161	1.0000

**Note:** The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

Table C6

Pipeline-to-ship bunkering ATT regression, without covariates.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Pipeline-to-ship LNG bunkering in port	-2.7306	1.2083	-2.2599	0.2469
B. smooth terms	edf	Ref.df	F-value	p-value
Port random effects (random intercept)	14.4501	16.0000	504907.2716	1.0000
Vessel random effects (random intercept)	9.4508	4.0000	717377.2920	1.0000

Note: The asterisks associated with the estimated coefficients represent the significance levels. Significance at the 10 %, 5 % and 1 % levels are marked by one, two, or three asterisks, respectively (\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ ). The contribution of the random effects follows the significance of the associated *effective degrees of freedom* (edf). Significant edf indicates a strong contribution of the random effects. The Ref.df term denotes the reference degrees of freedom.

## References

- AB Klaipėdos nafta website, n.d. Klaipėda LNG terminal - KN [WWW Document]. URL <https://www.kn.lt/en/lng-terminals/klaipeda-lng-terminal/559> (accessed 8.25.23).
- Acciaro, M., 2014. Real option analysis for environmental compliance: LNG and emission control areas. *Transp. Res. Part D: Transp. Environ. Emiss. Control Areas Impact Marit. Transp.* 28, 41–50. <https://doi.org/10.1016/j.trd.2013.12.007>.
- Accocella, A., Caplice, C., Sheffi, Y., 2020. Elephants or goldfish?: An empirical analysis of carrier reciprocity in dynamic freight markets. *Transp. Res. Part E: Logist. Transp. Rev.* 142, 102073 <https://doi.org/10.1016/j.tre.2020.102073>.
- Adams, M., Quinonez, P., Pallis, A.A., Wakeman, T., 2009. *Environmental issues in port competitiveness*. Atlantic Research Report, Gateway Research Initiative Working Paper 7.
- Aghion, P., Bénabou, R., Martin, R., Roulet, A., 2023. Environmental preferences and technological choices: is market competition clean or dirty? *Am. Econ. Rev.: Insights* 5, 1–20. <https://doi.org/10.1257/aeri.20210014>.
- Alamouh, A.S., Ölçer, A.I., Ballini, F., 2022. Ports' role in shipping decarbonisation: A common port incentive scheme for shipping greenhouse gas emissions reduction. *Clean. Logist. Supply Chain* 3, 100021. <https://doi.org/10.1016/j.clscn.2021.100021>.
- Anderson, C.M., Opaluch, J.J., Grigalunas, T.A., 2009. The demand for import services at US container ports. *Marit. Econ. Logist.* 11, 156–185. <https://doi.org/10.1057/mel.2009.4>.
- Angrist, J.D., Pischke, J.-S., 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Aronietis, R., Sys, C., van Hassel, E., Vanelslander, T., 2016. Forecasting port-level demand for LNG as a ship fuel: the case of the port of Antwerp. *J. Shipp. Trd.* 1, 2. <https://doi.org/10.1186/s41072-016-0007-1>.
- Arpino, B., Mealli, F., 2011. The specification of the propensity score in multilevel observational studies. *Comput. Stat. Data Anal.* 55, 1770–1780. <https://doi.org/10.1016/j.csda.2010.11.008>.
- Ashbeck, E.L., Bell, M.L., 2016. Single time point comparisons in longitudinal randomized controlled trials: power and bias in the presence of missing data. *BMC Med. Res. Method.* 16, 43. <https://doi.org/10.1186/s12874-016-0144-0>.
- Baron, D.P., 1995. Integrated strategy: Market and nonmarket components. *Calif. Manage. Rev.* 37, 47–65. <https://doi.org/10.2307/41165788>.
- Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J., Stürmer, T., 2006. Variable selection for propensity score models. *Am. J. Epidemiol.* 163, 1149–1156. <https://doi.org/10.1093/aje/kwj149>.
- Brozynski, M.T., Leibowicz, B.D., 2022. A multi-level optimization model of infrastructure-dependent technology adoption: Overcoming the chicken-and-egg problem. *Eur. J. Oper. Res.* 300, 755–770. <https://doi.org/10.1016/j.ejor.2021.10.026>.
- Caliendo, M., Kopeinig, S., 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. *J. Econ. Surv.* 22, 31–72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>.
- Chen, K.K., Chiu, R.-H., Chang, C.-T., 2017. Using beta regression to explore the relationship between service attributes and likelihood of customer retention for the container shipping industry. *Transp. Res. Part E: Logist. Transp. Rev.* 104, 1–16. <https://doi.org/10.1016/j.tre.2017.04.015>.
- D'Este, G.M., Meyrick, S., 1992. Carrier selection in a RO/RO ferry trade Part 1. Decision factors and attitudes. *Marit. Policy Manag.* 19, 115–126. <https://doi.org/10.1080/03088839200000019>.
- De Leeuw, J., Meijer, E., Goldstein, H., 2008. *Handbook of Multilevel Analysis*. Springer.
- Delacrétaz, N., Lanz, B., Van Dijk, J., 2020. The chicken or the egg: Technology adoption and network infrastructure in the market for electric vehicles. IRENE working paper.
- Eide, M.S., Endresen, Ø., Skjong, R., Longva, T., Alvik, S., 2009. Cost-effectiveness assessment of CO2 reducing measures in shipping. *Marit. Policy Manag.* 36, 367–384. <https://doi.org/10.1080/03088830903057031>.
- Englert, D., Losos, A., Raucci, C., Smith, T., 2021. *The Role of LNG in the Transition Toward Low- and Zero-Carbon Shipping*. World Bank, Washington, DC.
- European Commission, 2014a. LNG Bunkering Infrastructure Solution and Pilot actions for Ships operating on the Motorway of the Baltic Sea (No. 2012- EU-21009–M). European Commission.
- European Commission, 2014b. HEKLA – Helsingborg & Klaipeda LNG Infrastructure Facility Deployment (No. 2014- EU- TM- 0120- W). European Commission.
- Fan, L., Gu, B., Luo, M., 2020. A cost-benefit analysis of fuel-switching vs. hybrid scrubber installation: A container route through the Chinese SECA case. *Transp. Policy* 99, 336–344. <https://doi.org/10.1016/j.tranpol.2020.09.008>.
- Ferrari, S., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. *J. Appl. Stat.* 31, 799–815.
- Gao, C., McDonald, R., 2022. Shaping nascent industries: Innovation strategy and regulatory uncertainty in personal genomics. *Adm. Sci. Q.* 67, 915–967. <https://doi.org/10.1177/00018392221112641>.
- Gasum LNG, 2021. Terminal Rules for the Pori LNG Terminal.
- Gelman, A., Hill, J., 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Goldstein, H., 2011. *Multilevel Statistical Models*. John Wiley & Sons.
- Haralambides, H.E. (Ed.), 2015. *Port Management*. Palgrave Macmillan UK, London. 10.1057/9781137475770.
- Hastie, T., Tibshirani, R., 1987. Generalized additive models: some applications. *J. Am. Stat. Assoc.* 82, 371–386. <https://doi.org/10.2307/2289439>.
- Heckman, J.J., LaLonde, R.J., Smith, J.A., 1999. The economics and econometrics of active labor market programs. In: *Handbook of Labor Economics*. Elsevier, pp. 1865–2097.
- Helcom, 2019. *Alternative Fuels for Shipping in the Baltic Sea Region*. HELCOM Baltic Marine Environment Protection Commission.
- Ho, D., Imai, K., King, G., Stuart, E.A., 2011. MatchIt: Nonparametric preprocessing for parametric causal inference. *J. Stat. Softw.* 42, 1–28. <https://doi.org/10.18637/jss.v042.i08>.

- Huber, S., Klein, E., Moeller, K., Willmes, K., 2015. Comparing a single case to a control group – Applying linear mixed effects models to repeated measures data. *Cortex* 71, 148–159. <https://doi.org/10.1016/j.cortex.2015.06.020>.
- IMO, 2015. *Third IMO GHG Study 2014*. International Maritime Organization, London, UK.
- ITF - OECD., 2018. Reducing Shipping Greenhouse Gas Emissions: Lessons from Port-Based Incentives. International Transport Forum and Organisation for Economic Cooperation and Development, Paris.
- Jiannan, C., Feng, L., Zhongzhen, Y., 2020. Impacts of the choice habits of port users on the effects and efficiencies of port investment. *Transp. Policy* 99, 203–214. <https://doi.org/10.1016/j.tranpol.2020.07.021>.
- Kim, J., Seltzer, M., 2007. Causal Inference in Multilevel Settings in Which Selection Processes Vary across Schools. CSE Technical Report 708. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Kim, S., Chiang, B., 2014. Sustainability practices to achieve sustainability in international port operations. *J. Korea Port Econ. Assoc.* 30, 15–37.
- Kim, A.-R., Kwak, D.-W., Seo, Y.-J., 2021. Evaluation of liquefied natural gas bunkering port selection. *Int. J. Log. Res. Appl.* 24, 213–226. <https://doi.org/10.1080/13675567.2019.1642311>.
- King, G., Nielsen, R., 2019. Why propensity scores should not be used for matching. *Polit. Anal.* 27, 435–454. <https://doi.org/10.1017/pan.2019.11>.
- Lam, J.S.L., Chen, D., Cheng, F., Wong, K., 2011. Assessment of the competitiveness of ports as bunkering hubs: empirical studies on Singapore and Shanghai. *Transp. J.* 50, 176–203. <https://doi.org/10.5325/transportationj.50.2.0176>.
- Leibowicz, B.D., 2018. Policy recommendations for a transition to sustainable mobility based on historical diffusion dynamics of transport systems. *Energy Policy* 119, 357–366. <https://doi.org/10.1016/j.enpol.2018.04.066>.
- Li, X., Sun, Y., Wu, G.T., 2013. Cost of LNG-fueled container vessels. *AMR* 805–806, 1325–1329. <https://doi.org/10.4028/www.scientific.net/AMR.805-806.1325>.
- Lirn, T.C., Thanopoulou, H.A., Beynon, M.J., Beresford, A.K.C., 2004. An application of AHP on transshipment port selection: A global perspective. *Marit. Econ. Logist.* 6, 70–91. <https://doi.org/10.1057/palgrave.mel.9100093>.
- LNG in Baltic Sea Ports, 2014. LNG Handbook.
- Magala, M., Sammons, A., 2008. A new approach to port choice modelling. *Marit. Econ. Logist.* 10, 9–34. <https://doi.org/10.1057/palgrave.mel.9100189>.
- Mäkitie, T., Steen, M., Saether, E.A., Bjørgum, Ø., Poulsen, R.T., 2022. Norwegian ship-owners' adoption of alternative fuels. *Energy Policy* 163, 112869. <https://doi.org/10.1016/j.enpol.2022.112869>.
- Malchow, M.B., Kanafani, A., 2004. A disaggregate analysis of port selection. *Transp. Res. Part E: Logist. Transp. Rev., Shipp. Finan. Port Issues* 40, 317–337. <https://doi.org/10.1016/j.tre.2003.05.001>.
- Mampaey, 2023. Equipping the Cryogas Vysotsk Terminal. URL <https://www.mampaey.com/equipping-the-cryogas-vysotsk-terminal/> (accessed 9.1.23).
- Martinez-Pardo, A., Orro, A., Garcia-Alonso, L., 2020. Analysis of port choice: A methodological proposal adjusted with public data. *Transp. Res. A Policy Pract.* 136, 178–193. <https://doi.org/10.1016/j.tra.2020.03.031>.
- McKinlay, C.J., Turnock, S.R., Hudson, D.A., 2021. Route to zero emission shipping: Hydrogen, ammonia or methanol? *Int. J. Hydrogen Energy* 46, 28282–28297.
- Mersky, A.C., Sprei, F., Samaras, C., Qian, Z. (Sean), 2016. Effectiveness of incentives on electric vehicle adoption in Norway. *Transp. Res. Part D: Transp. Environ.* 46, 56–68. <https://doi.org/10.1016/j.trd.2016.03.011>.
- Murphy, P.R., Hall, P.K., 1995. The relative importance of cost and service in freight transportation choice before and after deregulation: An update. *Transp. J.* 35, 30–38.
- Nguyen, T.-L., Collins, G.S., Spence, J., Daurès, J.-P., Devereaux, P.J., Landais, P., Le Manach, Y., 2017. Double-adjustment in propensity score matching analysis: choosing a threshold for considering residual imbalance. *BMC Med. Res. Method.* 17, 78. <https://doi.org/10.1186/s12874-017-0338-0>.
- Nir, A.-S., Lin, K., Liang, G.-S., 2003. Port choice behaviour—from the perspective of the shipper. *Marit. Policy Manag.* 30, 165–173. <https://doi.org/10.1080/0308883032000069262>.
- Offshore Energy, 2015. Skangas opens Risavika LNG bunkering station. URL <https://www.offshore-energy.biz/skangas-opens-risavika-lng-bunkering-station/> (accessed 9.1.23).
- Parola, F., Risitano, M., Ferretti, M., Panetti, E., 2017. The drivers of port competitiveness: a critical review. *Transp. Rev.* 37, 116–138. <https://doi.org/10.1080/01441647.2016.1231232>.
- Peng, Y., Zhao, X., Zuo, T., Wang, W., Song, X., 2021. A systematic literature review on port LNG bunkering station. *Transp. Res. Part D: Transp. Environ.* 91, 102704. <https://doi.org/10.1016/j.trd.2021.102704>.
- Port of Gothenburg website, n.d. Liquefied Natural Gas [WWW Document]. URL <https://www.portofgothenburg.com/about-the-port/greener-transport/alternative-fuels/> (accessed 8.25.23).
- Port of Helsinki, 2017. Safety Manual on LNG Bunkering Procedures for the Port of Helsinki (No. RE20167730- 01- 00- D). Port of Helsinki & SSPA.
- Rashbash, J., Browne, W.J., 2008. Non-hierarchical multilevel models. *Handbook of Multilevel Analysis* 301–334.
- Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55.
- Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66, 688–701.
- Salmenkaita, J.-P., Salo, A., 2002. Rationales for government intervention in the commercialization of new technologies. *Tech. Anal. Strat. Manag.* 14, 183–200. <https://doi.org/10.1080/09537320220133857>.
- Sánchez, R.J., Hoffmann, J., Micco, A., Pizzolitto, G.V., Sgut, M., Wilmsmeier, G., 2003. Port efficiency and international trade: port efficiency as a determinant of maritime transport costs. *Marit. Econ. Logist.* 5, 199–218. <https://doi.org/10.1057/palgrave.mel.9100073>.
- Sánchez, R.J., Ng, A.K.Y., Garcia-Alonso, L., 2011. Port selection factors and attractiveness: the service providers' perspective. *Transp. J.* 50, 141–161. <https://doi.org/10.5325/transportationj.50.2.0141>.
- Schielzeth, H., Nakagawa, S., 2013. Nested by design: model fitting and interpretation in a mixed model era. *Methods Ecol. Evol.* 4, 14–24. <https://doi.org/10.1111/j.2041-210x.2012.00251.x>.
- Schulz, F., Rode, J., 2022. Public charging infrastructure and electric vehicles in Norway. *Energy Policy* 160, 112660. <https://doi.org/10.1016/j.enpol.2021.112660>.
- Serry, A., 2017. Development of liquefied natural gas facilities in the Baltic Sea ports: a geographical perspective. *Reg. Form. Devel. Stud.* 23. <https://doi.org/10.15181/rfds.v23i3.1591>.
- Smålandshamn AB, 2019. LNG Bunkering Regulations, Port of Oskarshamn.
- Smith, A.J., Todd, E.P., 2005. Does matching overcome LaLonde's critique of nonexperimental estimators? *J. Econometr. Exper. Non-Exper. Eval. Econ. Policy Models* 125, 305–353. <https://doi.org/10.1016/j.jeconom.2004.04.011>.
- Stein, M., Acciaro, M., 2020. Value creation through corporate sustainability in the port sector: A structured literature analysis. *Sustainability* 12, 5504. <https://doi.org/10.3390/su12145504>.
- Struben, J., Serman, J.D., 2008. Transition challenges for alternative fuel vehicle and transportation systems. *Environ. Plann. B. Plann. Des.* 35, 1070–1097.
- Tongzon, J.L., 2009. Port choice and freight forwarders. *Transp. Res. Part E: Logist. Transp. Rev.* 45, 186–195. <https://doi.org/10.1016/j.tre.2008.02.004>.
- Ugboma, C., Ugboma, O., Ogwude, I.C., 2006. An analytic hierarchy process (AHP) approach to port selection decisions – empirical evidence from Nigerian ports. *Marit Econ Logist* 8, 251–266. <https://doi.org/10.1057/palgrave.mel.9100160>.
- UNCTAD, 2020. Review of Maritime Transport 2020. United Nations, Geneva.
- Valionienė, E., Strakauskaitė, A., 2015. Evaluation of bunkering services attractiveness at Klaipėda seaport after establishment of LNG terminal. *J. Marit. Transp. Eng.* 32–39.
- Veldman, S., Garcia-Alonso, L., Vallejo-Pinto, J.Á., 2011. Determinants of container port choice in Spain. *Marit. Policy Manag.* 38, 509–522. <https://doi.org/10.1080/03088839.2011.597450>.
- Wang, S., Notteboom, T., 2014. The adoption of liquefied natural gas as a ship fuel: A systematic review of perspectives and challenges. *Transp. Rev.* 34, 749–774. <https://doi.org/10.1080/01441647.2014.981884>.
- Wang, S., Notteboom, T., 2015. The role of port authorities in the development of LNG bunkering facilities in North European ports. *WMU J. Marit. Affairs* 14, 61–92. <https://doi.org/10.1007/s13437-014-0074-9>.

- Wechsler, S., 2021. Maersk leapfrogs LNG to net zero marine fuels [WWW Document]. accessed 5.24.23 Argus Media. <https://www.argusmedia.com/en/news/2190195-maersk-leapfrogs-lng-to-net-zero-marine-fuels>.
- Wood, S.N., 2017. *Generalized Additive Models: An Introduction with R*, 2nd ed. Chapman and Hall/CRC, Boca Raton. 10.1201/9781315370279.
- Yeh, S., 2007. An empirical analysis on the adoption of alternative fuel vehicles: The case of natural gas vehicles. *Energy Policy* 35, 5865–5875. <https://doi.org/10.1016/j.enpol.2007.06.012>.
- Yu, Z., Guindani, M., Grieco, S.F., Chen, L., Holmes, T.C., Xu, X., 2022. Beyond t test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron* 110, 21–35.
- Yuen, K.F., Wang, X., Wong, Y.D., Zhou, Q., 2018. The effect of sustainable shipping practices on shippers' loyalty: The mediating role of perceived value, trust and transaction cost. *Transp. Res. Part E: Logist. Transp. Rev.* 116, 123–135. <https://doi.org/10.1016/j.tre.2018.06.002>.
- Zink, R., Valdes, J., Wuth, J., 2020. Prioritizing the chicken or egg? Electric vehicle purchase and charging infrastructure subsidies in Germany. *Pol. Govern.* 8, 185–198. <https://doi.org/10.17645/pag.v8i3.3025>.