LONDON'S GLOBAL UNIVERSITY

UCL ENERGY INSTITUTE

Revealed preferences for energy efficiency in the shipping markets

Prepared for Carbon War Room
August 2016

Authors

Vishnu Prakash, UCL Energy Institute
Dr Tristan Smith, UCL Energy Institute
Dr Nishatabbas Rehmatulla, UCL Energy Institute
James Mitchell, Carbon War Room
Professor Roar Adland, Department of Economics, Norwegian School of Economics (NHH)

Contact

If you have any queries related to this report, please get in touch.

James Mitchell
+44 1865 514214
Carbon War Room
jmitchell@carbonwarroom.com

Dr Tristan Smith
UCL Energy Institute
+44 203 108 5984
tristan.smith@ucl.ac.uk

About UCL Energy Institute

UCL Energy Institute delivers world-leading learning, research, and policy support on the challenges of climate change and energy security. Its approach blends expertise from across UCL, to make a truly interdisciplinary contribution to the development of a globally sustainable energy system.

The shipping group at UCL Energy Institute consists of researchers and PhD students, involved in a number of on-going projects funded through a mixture of research grants and consultancy vehicles (UMAS). The group undertakes research using models of the shipping system (GloTraM), shipping big data (including satellite Automatic Identification System data), and qualitative and social science analysis of the policy and commercial structure of the shipping system.

The shipping group's research activity is centred on understanding patterns of energy demand in shipping and how this knowledge can be applied to help shipping transition to a low carbon future. The group is world leading on two key areas: using big data to understand trends and drivers of shipping energy demand or emissions and using models to explore what-ifs for future markets and policies.

Acknowledgments

This research was funded by the Carbon War Room and their Stranded Assets programme. The authors would like to thank Kris Fumberger and Wayne Blumenthal of Rightship, and Dr Sophia Parker for reviewing the report. Whilst not a component of this project, the study builds on experience and knowledge gained during the Research Councils United Kingdom (RCUK) Energy programme and industry-funded Shipping in Changing Climates (SCC) project. The authors would like to thank MSI for co-funding the SCC project and the PhD studentship that has enabled this research.

Contents

List of figures

List of tables

1 Foreword

James Mitchell
Senior Associate, Carbon War Room

Since the monumental achievement of the Paris Agreement in late 2015, the International Maritime Organization (IMO) has come under increasing pressure to address the international shipping industry's 2.4% contribution towards global GHG emissions. However, there is a significant disconnect between projected 'current policy' emissions, which estimate that shipping could contribute up to 18% of global GHG emissions by 2050, and the sector-wide emissions reductions that would be necessary to contribute to a global target of 1.5°C or 2°C warming. Bridging this decarbonisation gap may expose the industry to new market and policy risks, and could result in the premature stranding of assets and capital. Given this, it is necessary for industry stakeholders with long-term vested interests in those assets, specifically ship financiers and owners, to understand the potential impacts of policy-induced stranded assets.

This report, commissioned by Carbon War Room and completed in close cooperation with University College London, is part two of three in our work on stranded assets in shipping. Part one identified major financial stakeholders' awareness of stranded asset risks as well as potential methods for understanding and mitigating those risks. This report, part two, tests whether technical vessel efficiency measures can be used to anticipate those risks by establishing some understanding of the role of vessel efficiency in competitiveness in past markets. It is, to our knowledge, the most comprehensive work of this type to date, but our hope is that it will catalyse further work as laid out in the body of this study. Part three will use the outcomes of this report to investigate the impacts of decarbonisation pathways on vessel valuation and earnings.

Stranded assets in shipping are those vessels that have suffered from unanticipated or premature write-downs, devaluations, or conversion to liabilities. Due to the highly cyclical nature of shipping markets, the industry is no stranger to stranded asset risks. However, addressing the industry's carbon emissions will provide new risk factors capable of stranding significant amounts of vessels. The stranded assets framework is ideal to anticipate such risks because it evaluates risks at the asset or project level, where much of shipping newbuild and second-hand investment still occurs.

This approach has been taken in other sectors by overlaying infrastructure stock with climate pledges or decarbonisation pathways. Those assets most 'at risk' are identified and vested stakeholders can mitigate material risks. For example, as coal-fired power generation is often the first policy target for national decarbonisation strategies, six of the largest financiers of coal-fired power stations[1] have announced policies to end investment in the least efficient power stations because they are the most vulnerable to a raft of policies—particularly GHG policies.

Identifying stranded asset risks in shipping is arguably more complicated than in power generation, where assets can't relocate to find more favourable regulatory regimes or markets. The findings of this report confirm this; they suggest that more work is necessary to understand future risks and they suggest that anticipating those risks will be incredibly difficult if present market failures persist. Nevertheless, material risks loom for owners and financiers and must be understood.

---

[1] BNP Paribas, Credit Agricole, HSBC, PNC Financial, RBS, and Societe Generale - *The End of Coal? Coal Finance Report Card 2015*

Because shipping falls outside of the Paris Agreement, we must look to the IMO and other regulatory bodies for potential upcoming policies capable of stranding assets. While progress is admittedly slow, such risks should be considered as material because policy changes can realistically occur within the horizon of investment and ownership decisions made today. The best illustration of this is the Republic of the Marshall Islands' proposal to the Marine Environment Protection Committee's 68th session to limit GHG emissions of the sector in line with a 1.5°C expected rise in global temperatures. While this proposal was unsuccessful, it was supported by 47% of member states[2] at the plenary session where the proposal was discussed.

Tools that assess vessels on their design efficiency—based on the IMO's EEDI and Rightship's EVDI—are available to institutions to establish where their vessels stand relative to peer vessels. The incorporation of such tools should be commended and it is our view that they should sit alongside existing decision-making practices, which are specific to each institution and market. However, as the results of this study suggest, such tools may not provide the necessary insight to understand stranded asset risks in shipping, as they might in other industries. Part three of our work on stranded assets in shipping, which will be released later this year, will begin to fill in some of these gaps by investigating the impact of technically feasible decarbonisation pathways on vessel earnings and valuation.

This report raises many challenges for the shipping industry. We welcome further development and refinement of existing efficiency tools, the development of new tools that account for operational efficiency, and development of tools to anticipate climate-related risk. We would also challenge the industry and its financiers to push for policy certainty, to accelerate work on identifying 1.5 and 2°C pathways, and to understand the risks and benefits of these pathways. It is possible to decarbonise without unnecessary crisis, but there is far to go at present.

---

[2] MEPC 68 Agenda item 5, Reduction of GHG Emissions from ships: An analysis of the plenary session. See: www.u-mas.co.uk/Latest/Post/302

2 Executive summary

This study assesses whether there is sufficient evidence to suggest that the revenue and operational performance of vessels in the dry bulk and tanker markets are influenced by their greenhouse-gas (GHG) emissions ratings as a proxy for energy efficiency. That is, whether vessels with good emissions ratings, first, generate higher revenues and, second, are better or more extensively utilised than their peers with poor ratings.

Carbon War Room and Rightship's GHG Emissions Rating[3] is used throughout this paper[4] to characterise the technical energy efficiency of individual ships. The GHG Emissions Rating assigns a score between A and G using the vessel's estimated CO_2 emissions per tonne nautical mile relative to the vessel's peer group[5] (Rightship, 2015a). To overcome the issue of having too few vessels in some of the 7 individual GHG Rating[6] classes, vessels were grouped into 3 rating categories: A to C, D, and E to G, which splits the distribution of vessels in any of the fleets studied reasonably evenly. It is recognised that whilst important questions about the reliability of these ratings remain, they are increasingly used in commercial transactions and therefore can be expected to have some influence in the market.

We meet our first objective by estimating whether vessels with good GHG Ratings were earning higher time charter rates than those with poor GHG Ratings using fixtures data from 2005 to 2015 for the Panamax (60,000–100,000 dwt) and Capesize (100,000+ dwt) dry bulk fleets and the Suezmax (120,000–200,000 dwt) and VLCC (200,000+ dwt) tanker fleets. This model effectively tests for the existence of a two-tier market for vessels based on their GHG Ratings, which is most likely to exist in the time charter market because the burden of fuel expenditure falls on the charterer rather than the ship owner.

Using a model based on Parker & Prakash (2016) and Adland et al. (2015), we account for vessel characteristics, local market conditions prior to and around the time of each fixture, the heterogeneity in time charter rates observed over time, space, and for different contract lengths, and biases within the sample of fixtures. If, after controlling for these factors, the difference in time charter rates between vessels with good and poor GHG Ratings is of a significant magnitude, it can be considered a price signal or additional[7] incentive to invest in vessels with good GHG Ratings.

Assessing differences in operational performance between vessels with good and poor GHG Ratings requires a different set of tools. We measure the operational performance of the Capesize and VLCC fleets from 2012 to 2015 at both an annual and voyage[8] level. Measures include everything from the ballast and laden speeds, time and distances spent laden and in ballast, capacity utilisation, to productivity[9]. These measurements are derived from Automatic Identification System (AIS) data, which provide information on each vessel's speed, position, draught, and a few other parameters at

---

[3] www.shippingefficiency.org

[4] A discussion on the choice to use GHG Emissions Ratings and its limitations can be found in section §4, along with some of the problems encountered in §7.3.3.

[5] A vessel's peer group is defined as the collection of vessels of the same type within ± 10% of the vessel's designed deadweight.

[6] GHG Emissions Rating and GHG Rating may be used interchangeably.

[7] This is in addition to the implicit saving in fuel costs.

[8] A voyage is defined as roughly a ballast leg followed by a laden leg, hence capturing the cost in time and distance required to reposition the vessel to load new cargo.

[9] We define productivity as the ratio of the actual tonne nautical miles of work completed to the potential tonne nautical miles of work it could have completed if the vessel had sailed at design speed, carried maximum cargo, and had minimum idle time.

high frequency. AIS data is assimilated to classify vessels according to their loading condition (using draught), estimate the tonnage of cargo, and identify voyages and routes. If vessels with good GHG Ratings are, for example, spending less time in ballast, carrying proportionately larger cargoes, sailing longer distances laden, or travelling at speeds greater than their peers with poor GHG Ratings, these could be implicit signals that operators were potentially cognisant of its GHG Rating or the energy efficiency it proxies.

It is hoped that the understanding derived here across multiple shipping sectors and across a diverse range of market conditions will help guide the potential responses of those vessels in future carbon-constrained markets, as well as provide examples of data, methods, and tools that could be developed further to provide insight to investment and operational decision making.

2.1 Key findings

1. For the tanker and dry bulk fleets studied, little to no evidence of a preference for ships with better GHG Ratings is detected in time charter rates, reflecting that a portion of the fuel saving associated with these ships is passed back to the owner. Between 2005 and 2015, the strongest preference detected is for ships in the Panamax dry bulk fleet, where ships with good GHG Ratings earned on average 2% more than their peers with poor GHG Ratings.
2. There is no evidence that the increased awareness over time of the GHG Emissions Rating measure of relative energy efficiency has resulted in an increasing influence on this preference.
3. Ships with higher GHG Ratings are not, on average, performing more work or work more quickly: no significant difference is observed in terms of productivity for ships with better GHG Ratings. This implies that utilisation (time spent loaded and number of loaded voyages, for example) is fairly consistent, regardless of the ship's GHG Rating.
4. In combination, these price and utilisation findings suggest that the time charter earnings of a ship in today's low freight rate markets will be similar regardless of its GHG Rating.
5. There are small differences in terms of operating speeds, where ships with good GHG ratings are, on average, being operated at slightly lower laden and ballast speeds than ships with poor GHG Ratings.
6. The results also show that there is significant operational variability (especially in speed) within any given GHG Rating category, but no clear explanation can be attributed to why ships in the same category are operated at significantly higher or lower speeds than the median or average. The fact that ships do travel faster (or slower) than the fleet-wide average speed is interesting, because it suggests that there are operators or charterers who would have benefitted more significantly from a vessel with a good GHG Rating than others.
7. The findings of this report are the result of a series of methods developed over a number of years. The methods used are thought to represent the state of the art to reveal preferences for GHG Ratings through price and operational signals. Nevertheless, there are a few areas where the analysis performed in this report could be improved. Data and method refinements could provide greater fidelity in representing the shipping markets studied.

2.2 Implications

The implications of our results can be separated into those attributable to the key agents or stakeholders in the shipping markets: charterers, ship owners, operators, financiers, policy makers, and researchers.

2.2.1 Charterers

1. For charterers, there is no strong evidence that ships with good GHG Ratings are getting the implied higher efficiency recognised in their day rate. This is to the charterer's benefit. As a charterer, if you are not already doing so, you should be looking to charter ships with good GHG Ratings because, all else being equal, you stand to save on fuel costs without an added cost or premium for the privilege.
2. However, in the longer run, this might also be to a charterer's detriment. Charterers are using a service (shipping) provided by ship owners who have a very marginal incentive to compete on efficiency, particularly in today's market. Until this failure is addressed, charterers will be vulnerable to paying more in overall costs (day rate plus fuel costs) in a higher market. That is, without addressing the lack of a decisive incentive to bring increasingly efficient vessels into service, there is a probability that charterers will face a high-cost market in the future that will be exacerbated by the lack of energy efficient ships.

2.2.2 Ship owners and operators

1. The marginal excess returns to ship owners found in the Panamax dry bulk sector for vessels with good GHG Ratings may not be encouraging in terms of justifying the investment case for a more efficient new build or the undertaking of retrofit effort, but work establishing investment costs is necessary to assess this conclusively.
2. Nevertheless, they do show opportunity, particularly if operating ships with a good GHG Rating (owned or time-chartered in) on the spot market for which any fuel saving should accrue to the owner or operator.
3. Part of the explanation for the findings could be to do with the quality of the information that is used to characterise a ship's GHG Rating and enable informed decisions by charterers. Contributing towards the development of better indices or data would help address this and ultimately improve the returns for ship owners with more energy efficient ships.

2.2.3 Financiers

1. The lack of strong preference signals on the basis of GHG Emissions Ratings impacts how those ratings might be used by financiers looking to assess risks from potentially carbon-constrained future shipping markets. Further, GHG Ratings do not provide an indication of exposure to other potential risks (technology obsolescence, for example). Together, this suggests that, in the short term—although there may be other benefits and the opportunity to influence wider behaviour—there may only be a modest portfolio-return incentive for financiers to make investment decisions using information on a ship's GHG Rating. Hence, a different and more sophisticated approach to assessing these risks may be necessary.
2. Nonetheless, there remains a probability that financiers willing to buck the trend and invest in energy efficient ships today may be rewarded in the longer run in a carbon-constrained scenario, if others failing to anticipate these rewards create a scarcity in energy efficient ships. There is therefore a clear need to develop better tools, create new measures, and explore simulations that can not only assess these types of scenarios, but also recommend best courses of action for financiers.

2.2.4 Policymakers

1. There remains evidence of market failures and barriers in shipping, even in spite of several years of availability and known use of GHG Ratings. Regardless of expectations of future fuel prices, this justifies the further use of regulation to assist in driving energy efficiency or decarbonisation. That regulation could include steps to address the market failures directly

(for example, by correcting information gaps and asymmetry), as well as a mandatory efficiency standard if that standard does not distort the market or induce unintended consequences.

2. The effectiveness of price mechanisms to enable GHG reduction (for instance, a carbon price or fuel levy) will only be weak at incentivising the design and purchase of more energy efficient ships. Without first addressing the market barriers and failures implied by the results in this report, extreme price signals or corrections may be required to achieve significant changes in vessel technology, operation, and GHG emissions.

2.2.5 Research

1. We have shown that tools that derive information from AIS data can provide large samples of results at a heretofore-unprecedented level of operational detail, which can help in the analysis of trends and behaviour in shipping markets.
2. There are multiple drivers of prices (day rates) in shipping markets, and isolating the influence of individual parameters (a ship's GHG Rating, for instance) is challenging but possible. Nonetheless, there is plenty of scope for the methods described in this report to be further refined.
3. Vessel speed remains a variable that is hard to fully explain or attribute, and significant variability exists within fleets of similar ship type, size, and technical specifications (including GHG Rating). Given speed's significance to operational energy efficiency, further work that examines the drivers of speeds will be important for understanding the sector's GHG emissions.

The results generated in this report highlight the many issues present in both the way that shipping markets function when pricing GHG Emissions Ratings into transactions and the apparent discord between operational performance and GHG Emissions Ratings. The intricacies of the market as well as the presence of information barriers may have obfuscated the signals we had hoped to detect, such that addressing these fundamental issues and barriers within the markets may increase the importance of measures of as-designed relative energy efficiency. Nevertheless, a more comprehensive solution is necessary to properly address stranding risks, and we hope to get a better picture of what such a solution might look like when potential decarbonisation pathways are identified and explored in the next part of our stranded assets work.

3 Introduction

Shipping faces multiple environment-related risk factors that can result in stranded assets, for example from changing technologies (the ability to bunker in an evolving fuels market or the ability to compete commercially against ships able to use cheaper fuels), energy efficiency (the ability to operate cost effectively and compete against more efficient designs), evolving vessel specifications (changes in size categories due to infrastructure evolution and canal constraint relaxations), to newly priced negative externalities (through carbon pricing or other forms of emissions regulation).

This paper focuses on the energy efficiency risk factor and assesses if there are signals in the market that indicate a preference for energy efficiency. Anecdotal evidence (Lloyds List 2011, 2012, 2013, Agnolucci et al., 2014, Smith et al., 2013) suggests the existence of a two-tier market between efficient and inefficient vessels, furthered by the increase in performance monitoring and publicly available data on energy efficiency. We examine this possibility by building on previous work analysing whether more energy efficient vessels attract higher prices in the market for shipping contracts (Smith et al. 2013, Agnolucci et al. 2014, and Adland et. al. 2015), but extend that to assess whether there are differing patterns of productivity and utilisation between energy efficient and inefficient vessels from a common fleet. Hence, under a hypothesis that, given a choice to utilise one of two vessels comparable in every aspect except energy efficiency, the more energy efficient vessel ought to always be preferable, this paper attempts to reveal an explicit preference for energy efficiency by looking for price signals in the markets for shipping contracts as well as an implicit preference for energy efficiency by examining vessel operational patterns.

Price signals are most likely to be detected in the time charter market where the burden of fuel costs falls on the charterer rather than the ship owner or operator. If the charterer here is privy to the differentiation in each vessel's energy efficiency, deems that information reliable, and anticipates the fuel savings accruing from the difference to be of sufficient value, the more energy efficient vessel ought to command a premium reflecting its cost saving potential relative to its less efficient alternative. The response of the time charter rate to changes in energy efficiency is estimated using the approach of Parker & Prakash (2016), a study similar in purpose to that of Agnolucci et al. (2014) and Adland et al. (2015) but different in methodology.

There may, however, be several factors that supersede energy efficiency or even negate the option to express a preference for energy efficiency. Excess supply or excess demand and prevailing fuel prices could temper both the preference for energy efficient vessels and the magnitude or need for a price premium. Indeed, a charterer looking for a vessel will be limited to an available subset that matches his requirements at that point in time. The proximity of the vessels to the hire area, as well as their technical specifications, costs, sanctions, and safety records could all potentially outweigh the importance of energy efficiency, such that the filtered subset of vessels may not proffer an opportunity to exercise a preference for energy efficiency. There may also be agents in the market who are unconcerned by energy efficiency altogether. A system with agents holding competing preferences could potentially mix or conceal the signals we wish to discover. Indeed, including the role of brokers or even the predispositions of charterers to vessels from certain owners could further complicate the signal. Nevertheless, it is hoped that, by estimating the model over a diverse range of market conditions, any consistent underlying pattern towards energy efficiency may be revealed. Employing Parker & Prakash's (2016) model allows us to search for such a price signal not only whilst controlling for localised market conditions and vessel specifications, but also whilst mitigating statistical issues present in the data.

An implicit preference for energy efficient vessels may exist irrespective of the contract or market in which the vessel operates. By exploiting satellite and terrestrial tracking data from the Automatic

Identification System (AIS) on vessel speeds, loading conditions, and trading patterns, the existence of an implicit preference is explored by looking at whether energy efficient vessels are better or more exhaustively utilised over a given period of time, within and between common trading routes, and across both the dry bulk and tanker markets. To our knowledge, this is the first study of its kind to estimate the relationship between observed, AIS-derived measures of utilisation and productivity to energy efficiency. We expect to observe that efficient vessels sail faster, spend less time in ballast, and are more productive compared to their nearest inefficient peers.

As with the attempt to detect price signals, market conditions and conflicting preferences could also affect the way that vessels are operated. Operational limitations enforced by external factors like weather and port congestion, as well as conditions on speed, fuel consumption, and penalties for delays embedded into the charter party could potentially disturb the expected relationship. It is hoped that looking at utilisation and productivity at an annual average level, on individual routes, and by exploring that behaviour over many years may reveal any underlying differences between the way efficient and inefficient vessels are operated.

Searching for these preferences for energy efficiency is motivated by the need to understand whether financial incentives exist for investing in energy efficient vessels or efficiency improving technology. Understanding what the relationships are between vessel earnings (or operational costs) and energy efficiency and how those relationships might respond to changes in the market could also facilitate studies of the implications of emissions related policies. Such a study could help identify vessels that are likely to become too costly to operate or become stranded as an asset if it becomes incapable of generating income in future, potentially carbon-constrained markets. Insights derived by studying revealed preferences could eventually be used to develop models that appraise vessel values over long horizons under policy-driven scenarios and abet potential investors in making better long-term decisions.

4 Measuring energy efficiency

Energy efficiency can be measured in a smorgasbord of ways (see Table 1, reproduced below from Smith et al., 2013). Operational energy efficiency that captures fuel consumption in recent and real-life conditions—across the vessel's spectrum of speeds—would appear ideal. However, it is often very difficult to measure and obtain actual fuel consumption in real world conditions (noon reports, for example, may not always be very reliable, Third IMO GHG Study, 2014), and, therefore, there is the potential for asymmetric information between the charterer and ship owner. Indeed, a failure to choose the right measure on our part could present another line of asymmetry.

The GHG Emissions Rating system developed by Rightship[10] based on EVDI (Existing Vessel Design Index) values that measure the grams of CO_2 emitted per tonne nautical mile—a measure of as-designed technical efficiency—is used throughout this paper. This system assigns a rating between A to G, and is designed to provide easy-to-interpret ratings of the relative energy efficiency of vessels. Charterers, banks, ship owners, and research institutions are known to utilise them, and public access is provided online[11]. Four factors led to our choice to use GHG Emissions Ratings as a proxy for energy efficiency.

First, EVDI and a GHG Rating based on EVDI is calculated and available for vessels built both prior to and after 2013 when the Energy Efficiency Design Index (EEDI, a measure equivalent in principle and in grams of CO_2 per tonne nautical mile) became mandatory for new vessels (IMO, 2010), making values available for the majority of active vessels in the fleets studied in this paper. Second, owners are permitted to update or correct EVDI values by providing certified vessel specification data or official EEDI scores, rendering Rightship's collection potentially more accurate than one that could be built manually from a registry of vessel technical specifications[12].

Third, data are available on the GHG Emissions Rating's increasing up-take, in contrast to the lack of data on the up-take or use of other measures. Total yearly requests by ship owners, charterers, researchers, and other stakeholders for GHG Ratings provided by Rightship have more than tripled since its introduction in late 2010 (from about 30,000 to close to 100,000 by the end of 2015). Further, Cargill, Huntsman, and UNIPEC UK publically announced in October 2012 that they would no longer charter the least efficient ships in the fleet (Rightship, 2012). Currently, 25% of the tanker and dry bulk market and 26 charterers use Rightship's GHG Emissions Rating classification as policy when choosing which vessels to charter (Lloyds List, 2015), and it is estimated that this amounts to annual movements for about 26,000 vessels and 2 billion deadweight tonnes of cargo (Rightship, 2015a).

Fourth, GHG Ratings that capture a vessel's EVDI or energy efficiency relative to its immediate peer group are perhaps suited to the way that charterers typically hire vessels. That is, a GHG Rating (between A and G) is applied to a vessel given its EVDI value and the EVDI values of vessels within a ±10% boundary around its deadweight (Rightship, 2015a, p. 13). Charterers should search for vessels that are within a narrow boundary of the size of the cargoes they expect to carry, which coincides with the way GHG Ratings are applied.

---

[10] www.rightship.com

[11] www.shippingefficiency.org

[12] About 23% of the 1158 oil tankers equal or above 120,000 tonnes studied in this paper utilise one or more owner corrected parameters or uses owner verified EEDI values. 47% of the fleet of 4518 dry bulk vessels of 60,000 tonnes or more have EVDI values that have been corrected in a similar way.

Term	*Description*	*Practical Considerations*
As-designed technical efficiency	The efficiency of a ship in its as-designed condition (straight from the yard) in ideal conditions.	This is what is captured in the EEDI (or EVDI) when it is applied to newbuild ships.
Technical efficiency in real operating conditions	The efficiency of a ship (straight from the yard) in real conditions (wind and waves etc.).	Careful attention to the hydrodynamics of a vessel in waves can save significant (20% and in some cases more) fuel consumption in actual use, but such benefits are not captured in the present EEDI formulation.
Technical efficiency at a point in time	The efficiency of a ship of a certain age, following wear, deterioration and fouling, benchmarked in ideal conditions	Heavy fouling can increase fuel consumption by up to 40-50% for a low speed ship (e.g. wet/dry bulk).
Measured technical efficiency	The efficiency of a ship of any age and condition, measured from fuel consumption but assuming 100% capacity utilization	Measurements of fuel consumption from trial specification (e.g. specified speed and draught) produce data on a ship‘s measured technical efficiency, which can in turn be validated e.g. by a classification society.
Transport supply efficiency	This embodies the relationship between the transport demand (e.g., tonnes of a commodity shipped), with actual capacity-distance (e.g., deadweight times nautical miles sailed)	Often, assumed 100% capacity utilization ignores the backhaul voyage emissions (regardless of vessel loading, ballast), which is virtually never the case.
Achieved operational efficiency	The energy consumed to satisfy a given transport demand	This could be considered the ultimate measurement of a ship‘s estimated real-world efficiency in that incorporates all of the components listed above.

Table 1: Some different definitions of energy efficiency (Smith et al., 2013, p. 5)

5 Price signals

One of the commonly cited market barriers to the uptake of energy efficient technologies or investment in green vessels in the shipping industry is the split incentive created in the time charter market (see Wilson, 2010, Wang et al., 2010, Faber, Behrends & Nelissen, 2011, and Rehmatulla, 2014), because voyage costs, including fuel costs, are borne by the charterer whilst capital and maintenance costs are borne by the owner or operator. The existence of a revenue incentive to invest in energy efficient ships depends on the degree to which owners in this market are rewarded through higher time charter rates and the proportion of time a ship spends in time charter.

By exploiting historical contract data from the time charter market, it may therefore be possible to understand whether such incentives exist, quantify the magnitudes of those incentives, and capture the responsiveness of those incentives to different market conditions. We do this for the Panamax (60,000–100,000 dwt) and Capesize (100,000+ dwt) dry bulk fleets, as well as the Suezmax (120,000–200,000 dwt) and VLCC (200,000+ dwt) tanker fleets.

5.1 Approach

Agnolucci et al. (2014) provide the first econometric attempt to this end, but are unable to establish the direct significance of a measure of technical energy efficiency on time charter rates. Using a benchmark vessel and building a model of relative fuel expenditure to that benchmark, they are instead able to show that only about 40% of the fuel savings from a more fuel-efficient ship is passed back to the owners through a higher time charter rate. Adland et al. (2015) find that energy efficiency as measured by EVDI is a significant—albeit marginal—determinant of time charter rates when estimated with dry bulk fixtures from 2001 to 2014. However, when testing the influence of *fuel* consumption per day and a fuel efficiency index (grams of fuel per tonne mile), they are unable to detect the existence of an explicit premium for fuel efficiency in any of the size categories in the dry bulk market. They caution the interpretation of the significance of EVDI, in part because the insignificance of their other measures of fuel consumption and efficiency raises the question of what EVDI might actually be measuring and in part because it would appear precocious to suggest that EVDI and energy efficiency measured in this way was used by the market before it was introduced in 2010 (IMO, 2010).

All of this, however, is completed, first, only with a sample of time charter *period* fixtures and excludes time-chartered trips and time-chartered round voyages, where, although the period of hire may be shorter, the appropriation of the fuel cost burden remains the same. Second, Adland et al. (2015) utilise EVDI explicitly, which, as an absolute measure, may be less effective than a relative measure if charterers hire vessels only from within a peer group of vessels best suited for the size of cargo they wish to move. Third, both Adland et al. (2015) and Agnolucci et al. (2013) do not account for the non-randomness of the dataset of fixtures obtained from a single source that is known to be biased to certain vessels and contract types, and neither tests the robustness of their models by applying it to a sector other than the dry bulk market.

Parker & Prakash's (2016) method is utilised in this report, which corrects many of the issues observed in the existing literature. Instead of absolute EVDI values, Rightship's relative GHG Emissions Ratings are used as a proxy for energy efficiency. Vessels are grouped into 3 rating categories (A-C, D, and E-G) to improve sample size and degrees of freedom issues that arise if using all 7 individual GHG Ratings in isolation. An estimate is derived of the percentage change in the time charter rate that is expected from moving from an A-C rated vessel to a D or an E-G rated vessel.

Unlike Adland et al. (2015) and Agnolucci et al. (2013), Parker & Prakash (2016) explicitly acknowledge that the distribution of time charter rates observed is conditional on, first, the publication of a fixture and its rate for a given vessel in the fixtures dataset, and, second, the allocation of that vessel to the time charter market rather than the spot market. It is very likely that any estimate of the effect of GHG Ratings on time charter rates derived using this conditional distribution will be biased unless the conditions that characterise the distribution have been controlled for. Econometric procedures employed in Parker & Prakash and adopted here temper these two sample selection problems.

This is done by estimating, first, the probability that at least one fixture is reported for any given vessel in the known fleet, and, second, the probability that at a particular point in time a vessel will be allocated to the time charter market rather than the spot market. Heckman's (1979) procedure is utilised to construct a covariate that corrects the bias captured by each probability model, but the bias-correction covariate from the first probability model is used in the second and the covariate from the second is used in the final model that estimates the influence of the GHG Rating category on the time charter rate—thereby accounting for both biases in the final model of the time charter rate (Behrman et al., 1980, Catsiapis & Robinson, 1982).

The first probability model is designed to account for biases in the sample attributable to the vessel's characteristics and the characteristics of its owner, which is necessary to control for issues like our data source not brokering fixtures for vessels owned by firms from certain countries or not fixing vessels beyond a certain age. The second probability model explains the likelihood of a vessel being allocated to the time charter market compared to the spot market as a function of vessel characteristics like its GHG Rating category, its historical predisposition to time charter contracts[13], its size class's historical predisposition[14], the income risk[15] in fixing on a long term time charter contract rather than on a spot contract, and the fuel cost risk[16] avoided by the ship owner in choosing to operate on the time charter market. This part is necessary because we essentially have a problem observed readily in models that estimate the determinants of wages. Wages are only observed for those who choose to work, and, hence, those who choose to work must be offered a wage that exceeds the value they place in leisure or not being employed. The distribution of wages we observe then is conditioned and affected by this choice. We believe the same principle applies here, given that an allocation to one of the two contract types takes place, such that being allocated to the time charter market defines the distribution of time charter rates observed.

5.2 Results

For the Capesize and Panamax fleet, the second probability model engenders a few interesting findings.

First, it is found that a vessel's and its fleet's historical predisposition to a particular contract provides the strongest determinant of the type of fixture the vessel will be allocated to at any point

---

[13] The ratio of the number of time charter fixtures to total fixtures for the vessel prior to current fixture is used to establish its historical predisposition to time charter contracts.

[14] Similar to the vessel's individual predisposition, this is measured by the ratio of the number of time charter fixtures to total fixtures for all vessels in the same size class and hired for the same on hire area in the 4 weeks prior to the current fixture.

[15] Measured as the volatility (standard deviation) in the ratio of the 6-month time charter rate to the time charter equivalent rate in the 12 weeks prior to the fixture. Average rates used are matched as closely as possible with the vessel's age and size class.

[16] The volatility in the average HFO price in the 4 weeks prior to the fixture.

in time. That is, a vessel that has historically been allocated to the time charter market is much more likely to continue operating on time charter contracts and is less likely to switch to spot contracts. This is also indicative of commercial rigidity in the dry bulk sector.

Second, we had hoped that the probability of being allocated to the time charter market would fall when income risk is high. That is, when the ratio of the rate received on a short time charter (6-month, in this case) to the time charter equivalent spot earnings is volatile before and close to the time of a given fixture, the likelihood of the vessel going into a time charter contract ought to fall. However, we do not find a significant relationship between income risk and the probability of a time charter.

We had also expected fuel price risk to exhibit a similar but opposite result: if the volatility in the fuel price around the time of the fixture was high, then the probability that the vessel enters the time charter market should increase. This would then have been perhaps indicative of a difference in the influence of the charterer and ship owner in the allocation decision, the presence of information asymmetry between the charterer and ship owner, or even a difference in the importance of fuel costs for the two agents under the prevailing market conditions. We are unable to find a significant relationship between fuel price risk and the allocation to the time charter market.

Third, we find no statistically significant indication that D or E-G rated vessels are more or less likely to be allocated to the time charter market than A-C rated vessels. That is, the GHG Rating of the vessel appears not to influence the contract type.

Estimates of the time charter *rate's* response to a change from an A-C rated vessel to a D or E-G rated vessel is then produced. In doing so, the model controls for the time period, seasonality, time charter type (trip, round voyage, 2-laden legs, and period), hire or delivery areas (large geographical blocks like North America or Europe[17]), and the bulker size classes included in the sample. With a statistically significant elasticity of 0.93, the most influential explanatory variable in the model is the average day rate received by vessels in the same size class hired for the same area and duration in the 4 weeks prior to the current fixture, supporting the findings in Adland et al. (2015). The hire area binary variables play an important role in capturing consistent spatial heterogeneity in the day rates, whilst the bulker size class dummy variables accentuate consistent size-specific differences in the day rates due to differences in fixed costs.

We find a small yet statistically significant premium for A-C rated Panamax vessels compared to both D and E-G rated vessels—confirming, to some degree, the marginal but significant influence detected in Adland et al. (2015). An A-C rated vessel returned on average a time charter day rate that is 2% higher than a D-rated vessel between 2005 and 2015. An A-C rated vessel earns a similar premium over E-G rated vessels, such that there appears not to be a sufficiently strong difference in the premium between D and E-G rated vessels.

There is no excess premium (or discount) for the Capesize fleet. This may be in part due to an insufficient sample of fixtures across a sufficiently diverse group of vessels within the Capesize fleet—compared to the Panamax subset—and in part due to the lack of variability in the gCO_2 per tonne nautical mile values (upon on which GHG Ratings are based) for the Capesize vessels in the sample of fixtures. A similar result was observed in Adland et al. (2015).

Whilst we were able to estimate how the premium varied across the two size classes in our sample, we were unable to establish how the *percentage* premium might have varied by year and under

---

[17] These capture consistent differences expected in rates as shown in Adland et al. (2015) where vessels delivered in the Atlantic have a premium to those delivered in the Pacific (China, in particular).

differing market conditions. Fracturing the data into year and rating category pairs did not always return a sufficiently large and diverse sample in each pair to produce consistent and statistically significant estimates of the premium (or discount).

Although the time charter market in the tanker sector is significantly smaller (Stopford, 2009), an attempt was made to extend the approach applied here for the dry bulkers to oil tankers (VLCC and Suezmax) over the same time period. However, similar results (including the existence of a premium) could not be established to an acceptable level of statistical certainty for the Suezmax and VLCC fleets.

Time charter fixtures formed a very, very small proportion (less than 5%) of all fixtures available to us for these two tanker size classes between 2005 and 2015, such that, first, our model estimating the likelihood of the allocation of a vessel to the time charter market did not perform as well as it did for the dry bulk fleet, and, second, left insufficient observations of fixtures across time and in each of the 3 GHG Rating categories to establish the presence and magnitude of any consistent relationship between the GHG Rating category and the time charter rate.

It remains to be tested whether, under the right market conditions, tankers with good emissions ratings are chartered in the spot market at a discount due to their implied lower fuel cost per mile. That is, our inability to establish the influence of GHG Ratings on the tanker fleet may not necessarily imply that such an influence does not exist, but suggests that we might simply require a different approach to unveil the influence.

6 Operational patterns

For Capesize bulkers and VLCCs, AIS-observed behaviour between 2012 and 2015 is used to ascertain whether the operational patterns of vessels differ significantly across the 3 GHG Rating categories.

At a yearly aggregate level, we investigate whether A-C rated vessels spend less time in ballast, carry proportionately larger cargoes, sail longer distances laden, or operate at speeds greater than their D or E-G rated peers—all of which could be indicators of the superiority in performance of vessels with good GHG Ratings and implicitly an indicator of the importance or influence of GHG Ratings in the markets (§6.2).

The differences are explored further in §6.3 at a higher resolution by comparing a similar set of measures of utilisation and productivity between vessels completing similar voyages.

6.1 Method

AIS messages emitted by a vessel provide a history of the vessel's progression in time and space, and changes in a vessel's reported draught and speed can be used to identify, for example, when it loads or unloads cargo or when it stops and starts sailing. However, the reliability of reported draught values tends to vary and draughts are often updated at times that do not concur with the arrival or departure from a port, which makes it difficult to identify both the event that took place at a port and its condition whilst sailing before or after a port call (Jia et al., 2015). Manual entry of draughts plays a part here, but so does intermittent loss in coverage from AIS receivers and censoring on the part of ship operators, governing entities, or both.

A set of verification procedures is thus introduced to minimise some of these uncertainties, albeit whilst ensuring that a sufficiently large, diverse, and unbiased sample of vessels in both time and space is retained.

For each vessel in a given year, an attempt is first made to identify its sequence of port calls. Every pair of port calls in the sequence is subject to strict verification conditions and is excluded from the construction of measures of the vessel's operational performance if verification fails. For instance, the journey between a given pair of port calls is excluded if the AIS-derived distance sailed falls outside of a tight window about the centre of the distribution of observed sailing distances for all vessels sailing between the same pair of ports.

Draughts reported closest to the point of arrival and after departure from a port are used to identify the event that occurred at that call. The entire sample of draughts reported by the vessel in its observed history is used to construct a bimodal distribution, typically with one mode occurring at the draught at which the vessel is in ballast and the other at the draught at which it is laden. The trough between these peaks is used to identify the draught at which the vessel's condition can be changed from ballast to laden (Smith et al., 2015, pp. 13-15).

Import and export data[18] by country for iron ore, coal, and crude oil further help verify these classified loading and discharging events. That is, a load or unload flag allocated to a port call is removed if the country at which the vessel stopped never exported or imported one or more of the commodities specific to that vessel type. Whilst the import and export data is not specific to

---

[18] These datasets were obtained from the United Nation's COMTRADE database.

seaborne trade nor to the size of the vessel, this step appears to correct port calls that are misclassified as loading or unloading by draught messages misaligned in time.

Given a vessel's sequence of port calls, these assignments made on its loading condition permit the splitting of the time, distance, and speed over which the vessel were operated into ballast and laden conditions[19]. Speeds between a pair or port calls is time-weighted to provide a more accurate representation of the journey, provided that messages are not too sparse and are not missing for significant portions of time[20]. These observations are then aggregated to provide a summary of each vessel's performance in a particular year. Vessels are only kept in the sample if at least 70% of the year passes verification and its observed behaviour is classifiable into both ballast and laden conditions. This does not, however, guarantee that all classifications applied across a vessel's history are unconditionally correct[21].

Whilst aggregating the behaviour of a vessel over each year provides indicators on its overall performance, a deeper examination is also permitted if the sequences of identified port calls and loading or unloading events can be chained together to form voyages. To allow an assessment of both the ballast and laden performance for any given voyage, a voyage is defined as a sequence of port calls in which it performs a ballast or repositioning leg followed by a laden leg. This structure accounts for the time and distance sailed after unloading a cargo, to loading a new cargo, and then sailing and unloading this new cargo. A vessel is assumed in ballast until its draught changes sufficiently to be classed as laden and, likewise, it is assumed laden until its draught changes sufficiently to be classed as in ballast, such that part loading and part discharging events are properly accounted for.

6.1.1 Inherited measurement errors

Many of the measures utilised here are reliant on draughts reported in AIS—from establishing the laden or ballast condition of the vessel at a point in time to estimating the cargo on-board.

Jia et al. (2015) construct a regression that estimates the cargo on-board from reported draughts for the Capesize fleet in 2012 by matching them to actual cargo weights reported by port agents in port call reports. In doing so, they find that nearly two thirds of the cargoes and loading or discharging events reported could be verified using AIS and draughts. Their regression is utilised in this paper to estimate the tonnage of cargo and compute a vessel's capacity utilisation. Although the regression is known to be very accurate for estimating cargoes for Capesize vessels from 100,000 to 250,000 tonnes, its efficacy wanes slightly beyond the 250,000 tonne mark (see Figure 3 and Table 8 in Jia et al., 2015). For oil tankers, the lightweight model developed in Smith et al. (2015, p. 16) based on Kristensen's (2012, 2013) work is utilised, which estimates payload by translating draught into total displacement and excluding a portion of that displacement due to the vessel's lightweight (the weight of its non-cargo components).

Distance travelled between two ports is computed as the sum of the distance between coordinates reported in every pair of AIS messages emitted between the two ports, such that, if messages are missing or are sparsely populated, the calculated distance will tend to diverge away from the true

---

[19] Implicitly, the method captures port time at each call, but this at present cannot be accurately classified between loading or discharging time and anchorage or waiting time.

[20] See §6.1.1 for a further discussion of potential measurement errors.

[21] If a vessel only reported a ballast draught over its entire AIS history, then the methods would assume that that draught was valid, and, in turn, measures using this would suggest that the vessel was unproductive. Filters are put in place to try to exclude such extreme cases.

distance sailed. No direct test is placed on each distance drawn between a consecutive pair of messages to ensure that the line between them does not cross land, but verification procedures are put in place to try to ensure that large deviations from actual distances between ports are flagged and excluded appropriately. The use and interpretation of measures of activity based on distance ought to nevertheless be cautioned.

Sparse AIS message streams also have implications on the accuracy of the time-weighted average constructed for speed between a pair of port calls. Speed reported in a message is assumed to remain constant until the next message is observed. Hence, if the next message is received only after a long period of time, the time spent at the last known speed may be exaggerated in the time-weighted representation if messages were missing unexpectedly. Sparseness of messages between a pair of port calls are nevertheless measured and checks are in place to detect and properly treat cases where extreme time-weighted speeds are computed.

6.2 Annualised performance

Each vessel's annual average capacity utilisation (the proportion of total deadweight filled with cargo), its average ballast and laden speeds, its laden and ballast distance ratios (laden or ballast distance over total distance sailed), and its laden and ballast time ratios (days laden or ballast over total days observed) are plotted as distributions over time. Another way to look at differences in speed is to consider ballast or laden speed a ratio of the vessel's design speed. Taking this ratio away from 1 gives a measure of how far from design speeds the vessels were actually being operated. We call this the "slack" in speed. A high % of slack in speed means the vessel was being operated further away from its design speed.

The distributions are split by year and the 3 GHG Rating categories (A-C, D, and E-G), and doing so permit a visual inspection of the shifts in the mean, median, standard deviation, the extent of dispersion within each subsample over time and, importantly, between the 3 GHG Rating classes. Observations in the distribution plots further show the extent of diversity in EVDI within each GHG Rating category via each observation's colour, where very similar colours indicate a lack of diversity[22]. Explicit values for the illustrated moments and measures of dispersion for these distributions can be found in Table 4 (Capesize) and Table 5 (VLCC).

The statistical significance of any observed variation illustrated in the distributions is measured by applying Wilcoxon's rank sum test (Wilcoxon, 1945) on the equality of medians. That is, an application of this test helps ascertain whether a given pair of medians from the separated samples contains values significantly different from each other by evaluating the null hypothesis that the medians are equal against the alternative that they are not. Non-normality, small sample sizes, and the presence of outliers necessitate a test like Wilcoxon's that is more robust to these issues (Wilcoxon, 1945). Results of these tests can be found in Table 6 and Table 7.

6.2.1 Capesize

Consistent, statistically significant variations in the results observed for Capesize bulkers occur predominantly in the distributions of average ballast and laden speeds between the 3 GHG Rating categories. Here, median ballast speed in the A-C class is about 0.2 knots slower than the median in E-G, and the median in D is 0.07 knots slower than E-G. Laden speeds are slower by 0.3 knots

---

[22] Some variation is expected, given the way in which GHG Ratings are assigned (Rightship, 2015a).

between A-C and E-G and by 0.2 knots between D and E-G. When broken down by year, the largest difference between two rating classes occur in 2012 between A-C and E-G, where A-C rated vessels were sailing about 0.7 knots slower than their E-G rated peers.

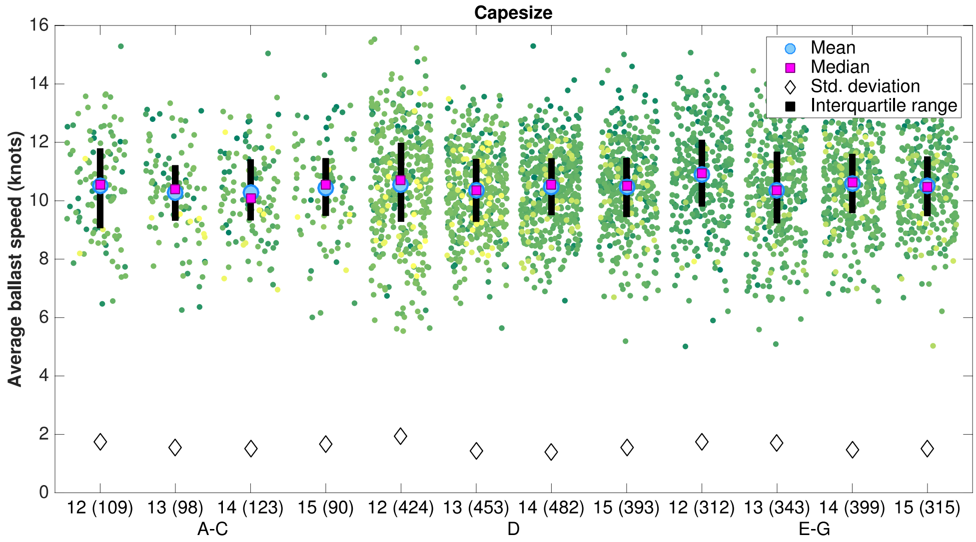

Figure 1: Annual average ballast speeds for the Capesize fleet

The differences are marginal and slower ballast speeds in A-C rated vessels, for example, have not translated to a difference large in magnitude for total time spent in ballast relative to either class D or E-G rated vessels, as shown in Figure 2. Large drops in the median can be found within the same rating class over time. For example, the ballast ratio for E-G rated vessels falls from 0.41 to 0.34 between 2012 and 2013—a 20% drop, implying a reduction in ballast time of around 50 to 70 days in 2013 (see Figure 3).

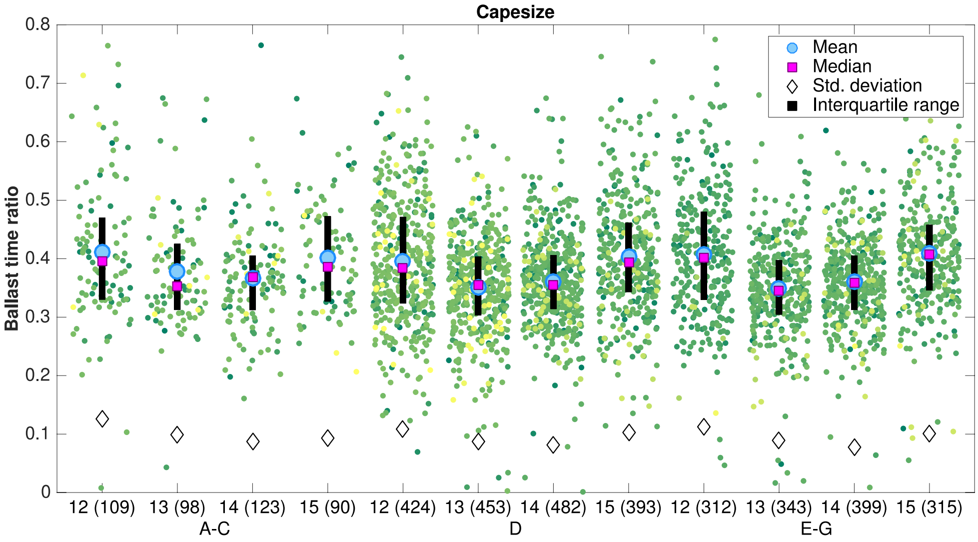

Figure 2: Ballast time ratio for the Capesize fleet

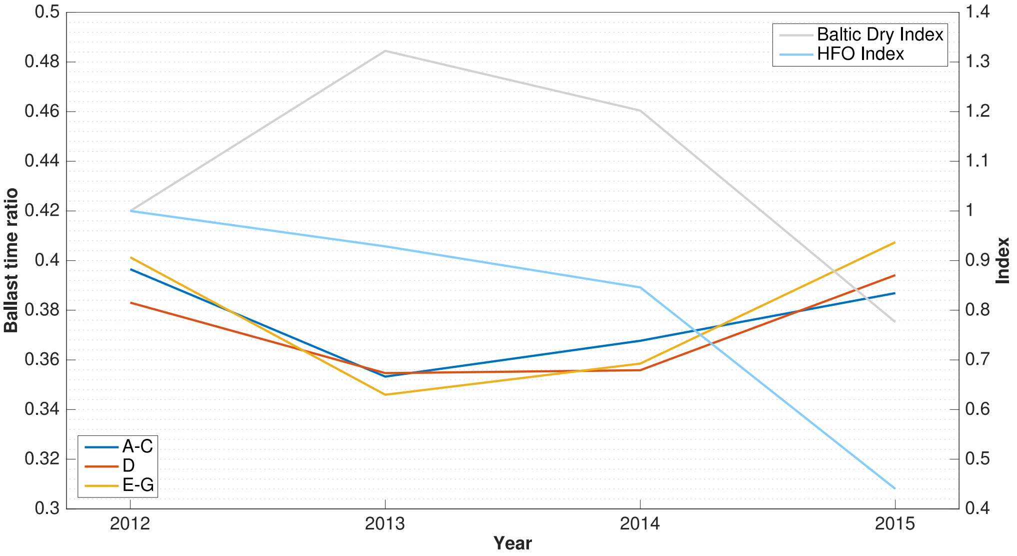

Figure 3: Median ballast ratios over time for the Capesize fleet

The medians in Figure 3 could, in general, be argued to be relatively stable across the three classes, more so in A-C and D than in E-G. It also suggests that the time spent in ballast as a proportion of total time observed is more strongly correlated to the prevailing rates as represented by the Baltic Dry Index than it is to fuel prices represented by the HFO Index, though this might be discouraging from an energy efficiency or emissions standpoint.

A pattern in average annual laden speeds similar to that found in Figure 1 is evident, where the medians suggest that A-C vessels are operated when laden at slightly slower speeds than D or E-G rated vessels.

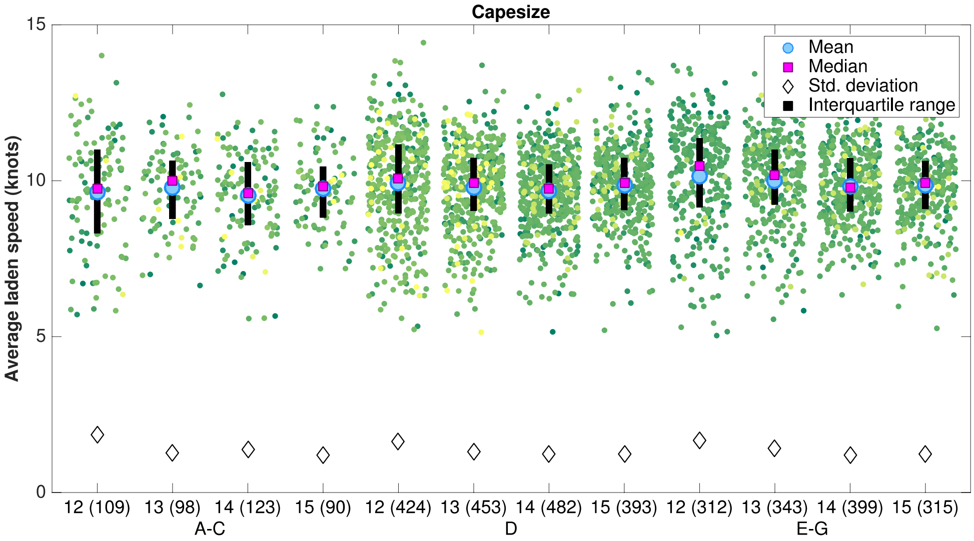

Figure 4: Annual average laden speeds for the Capesize fleet

The differences are again marginal, however. A-C rated vessels travel at about 0.3 knots slower compared to E-G, and D-rated vessels 0.2 knots slower than E-G. The differences we observe between absolute average speeds translate into significant differences in slack. For ballast speeds, we find that A-C vessels have 2% higher slack at 29% than E-G rated vessels, and D-rated vessels are about 1% higher than E-G. A-C rated vessels have a laden speed slack that is about 2.3% higher than E-G, 1% higher compared to D, and the difference between D and E-G at about 1.4%.

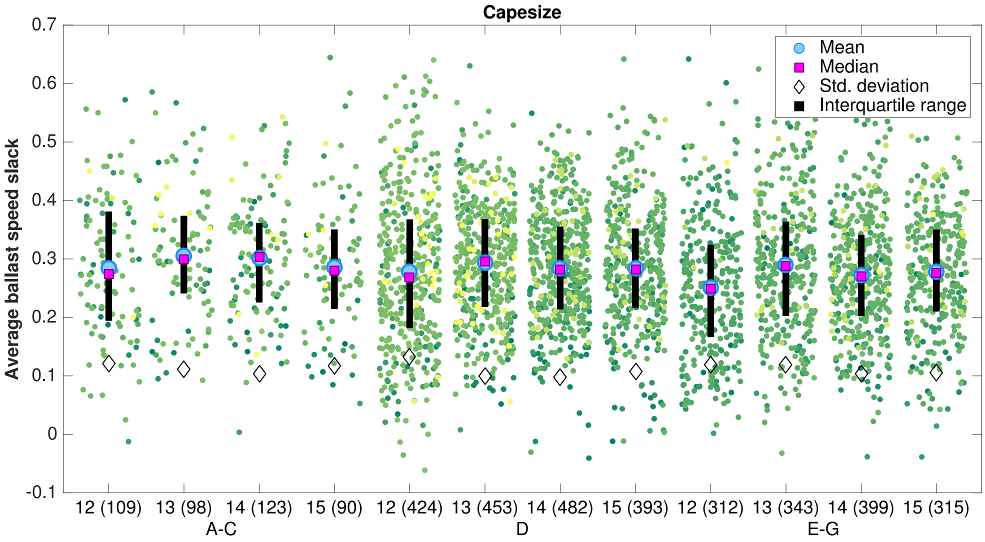

Figure 5: Slack in ballast speed

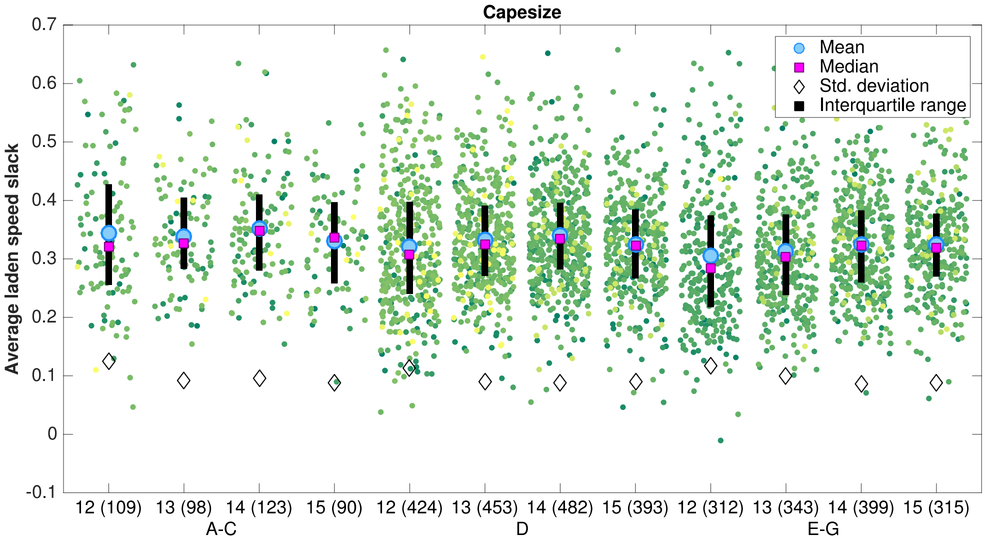

Figure 6: Slack in laden speed

We observe slight differences in capacity utilisation, where, for example, A-C rated vessels have lower capacity utilisation by about 0.8% compared to E-G. However, this difference is within the margin of error of the methods used to calculate cargo from draught (§6.1.1).

There are also differences in the measures in specific years: for instance, we find that the largest difference in ballast speed occurs in 2014 between A-C (10.1 knots) and E-G (10.6 knots) rated vessels. There are no consistent patterns in such differences with year-by-year analysis for other measures of operational behaviour.

6.2.2 VLCC

If all years were pooled together and Wilcoxon's rank-sum tests were applied, statistically significant differences in the median are detected only for ballast speed, laden speed, and capacity utilisation.

The median annual average ballast speed for A-C rated vessels is lower by 0.6 knots compared to E-G, laden speed is lower by 0.8 knots, and capacity utilisation is lower by just under 2%.

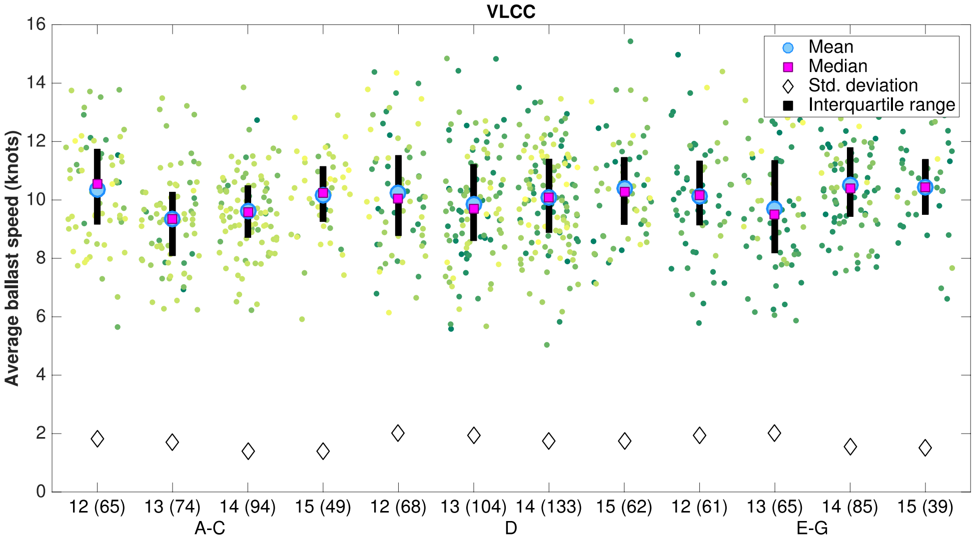

Figure 7: Annual average ballast speeds for the VLCC fleet

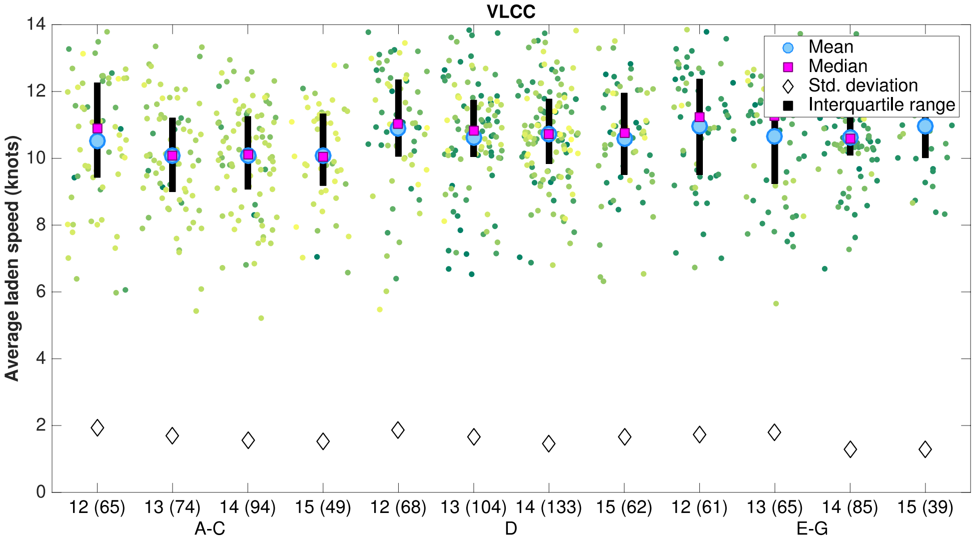

Figure 8: Annual average laden speeds for the VLCC fleet

The difference in magnitude for capacity utilisation is minor and is well within the margin of error of the method used to estimate cargo on board for VLCCs (Smith et al., 2015).

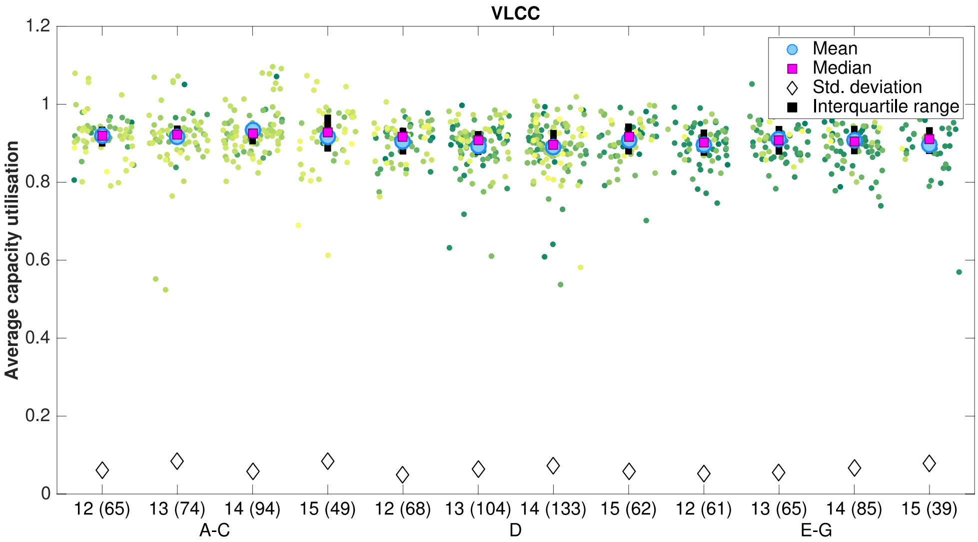

Figure 9: Annual average capacity utilisation in the VLCC fleet

No significant differences could be detected between the medians of the 3 GHG Rating classes for the vessel's ballast time ratio or laden time ratios. That is, there is no evidence to suggest that VLCCs with good GHG Ratings were spending more or less time in ballast or laden.

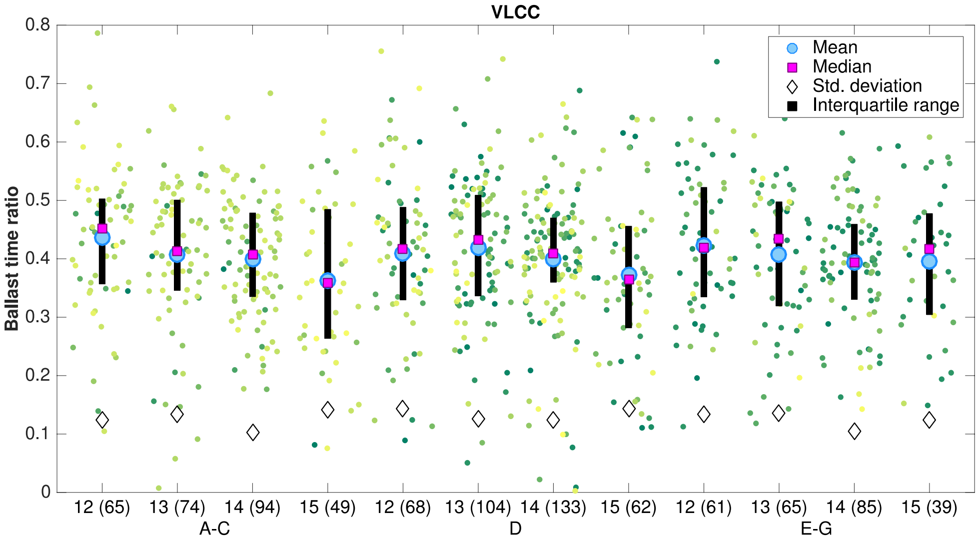

Figure 10: Ballast time ratios for the VLCC fleet

There are a few issues plaguing results for VLCCs. One is the lack of a sufficiently large sample of vessels, which is aggravated when the sample is split into GHG Rating classes and compromises the power of the rank-sum test used to detect differences in the medians. The other is more fundamental and is the lack of diversity in EVDI values for the fleet. The standard deviations of EVDI values within each GHG Rating class are 0.05, 0.04, and 0.09, respectively. With mean EVDI values of 2.4, 2.5, and 2.7, this translates into coefficients of variation of only 2.3, 1.8, and 3.3%. The average A-C rated vessel is therefore very similar in EVDI to a D-rated vessel, for example, which might be why there is no difference significant in magnitude between the two GHG Rating categories. Further, the coefficient of variation across the entire sample of vessels is only 5.6%. Nonetheless, the subset of the fleet captured here has a maximum EVDI of 3 $gCO_2/t.nm$ and a minimum of 2.15 $gCO_2/t.nm$, which equates roughly to a 40% difference in fuel consumption.

6.3 Voyage-level performance

To evaluate voyage level performance, we isolated the 10 most frequently sailed laden routes by each vessel type across all 4 years studied. For vessels sailing these routes, the measures computed heretofore at the annual level were recalculated and a similar set of statistical tests were performed. The objective here is twofold.

First, it is hoped that a voyage level analysis could help explain why the extent of variation observed at the annual level in any given year was beyond what was possible purely from differences in technical specifications or by the margins of error introduced through the methods used to compute the yearly measures of speed, capacity utilisation, and other measures.

Second, comparing operational performance from a GHG Rating standpoint for vessels completing identical or near-identical voyages removes some of the external conditions we were unable to control for before and moves the comparison closer to like-for-like.

6.3.1 Capesize

A new measure added to the voyage level analysis is that of productivity, defined as the ratio of the actual productivity (tonne nautical miles of work completed divided by the time it took for the whole voyage, including the ballast leg) to the potential or theoretical productivity (the amount of work it could have completed had it sailed at full speed, carried maximum cargo, and had minimum idle time). Defined in this way, productivity could be considered a composite measure of the vessel's performance.

For Capesize bulkers, there are consistent differences in productivity across routes. In these plots, the laden leg is described on the *x*-axis, the number of voyages observed for each route is in parentheses, and the dots in the plot are coloured by the EVDI value of the vessel completing each voyage. Readily noticeable are differences in EVDI across routes, where voyages taking cargo from Indonesia to China appear to be executed by vessels with high EVDI values. Vessels with low EVDI appear to operate on the Brazil to China route, consistent with both the trading patterns for Valemaxes and the observation that they tend to have low EVDI scores because of their size.

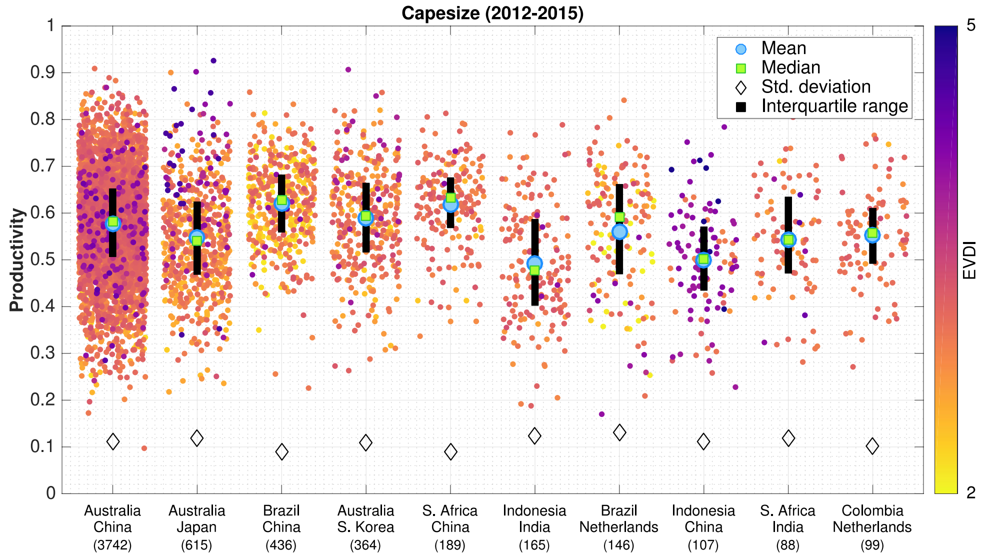

Figure 11: Productivity of the Capesize fleet by route

Whilst productivity appears to differ across routes, ballast time ratios, ballast speeds, and laden speeds display less variation across routes. This is indicative that most voyages for the Capesize fleet are loops or repeats with little opportunity to triangulate.

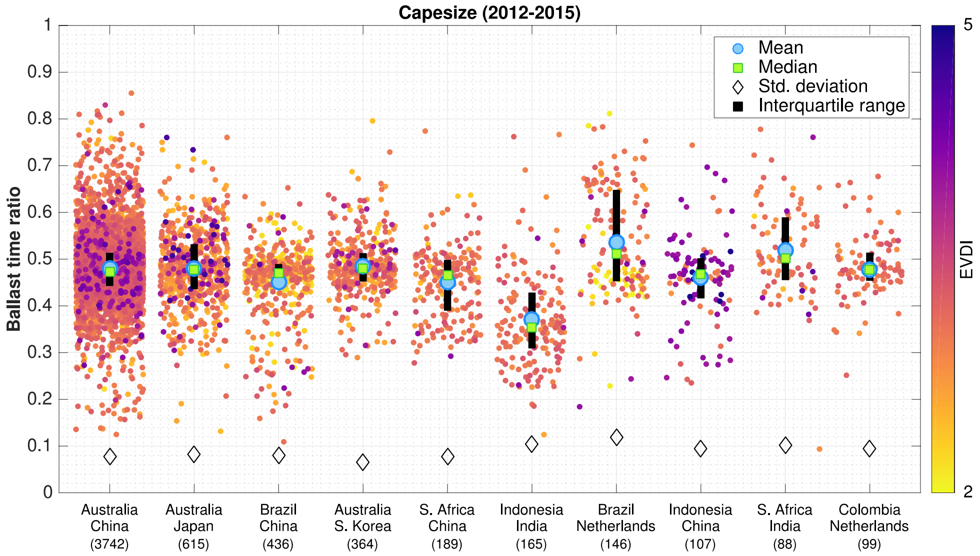

Figure 12: Ballast time ratios for the Capesize fleet by route

For speeds, we notice that at the voyage level, speeds revert to what they ought to be and display fewer effects of averaging seen at the annual level. Although there is significant variation in speed, median speeds across routes appear to be very similar. As expected, prime tonnage used for the Australia to Japan trade appears to operate slightly faster in ballast and whilst laden at both the mean and median.

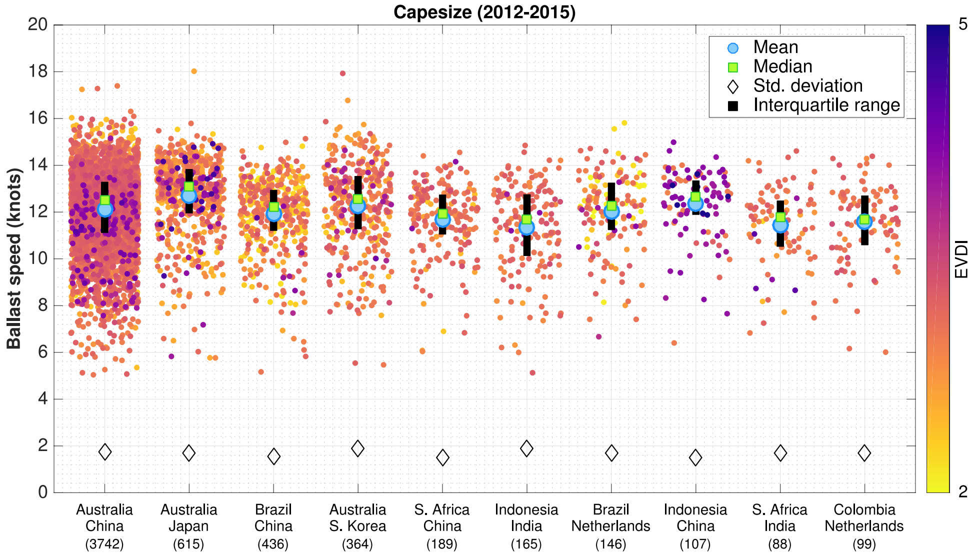

Figure 13: Ballast speed for the Capesize fleet by route

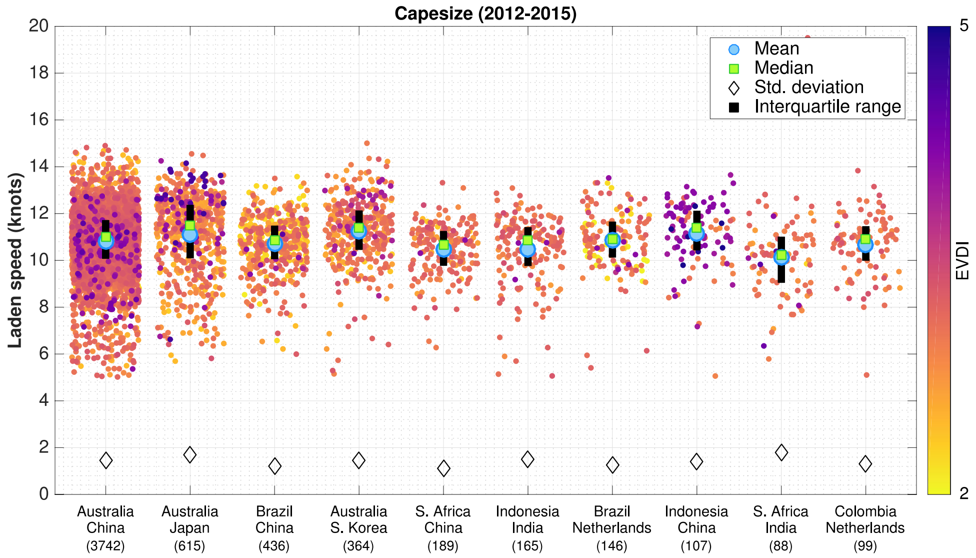

Figure 14: Laden speed for the Capesize fleet by route

Focusing on a particular route and splitting the voyages by GHG Rating categories permits an exploration of the influence of GHG Ratings across a similar set of observations. Consider the most frequent route between Australia and China. What is interesting immediately here is the influence of the route-level allocation of vessels seen at the EVDI level earlier, such that we can now see that fewer voyages by A-C rated vessels are completed on this route compared to voyages by D or E-G rated vessels. This might be indicative of a weak preference for vessels with good GHG Ratings from Chinese importers.

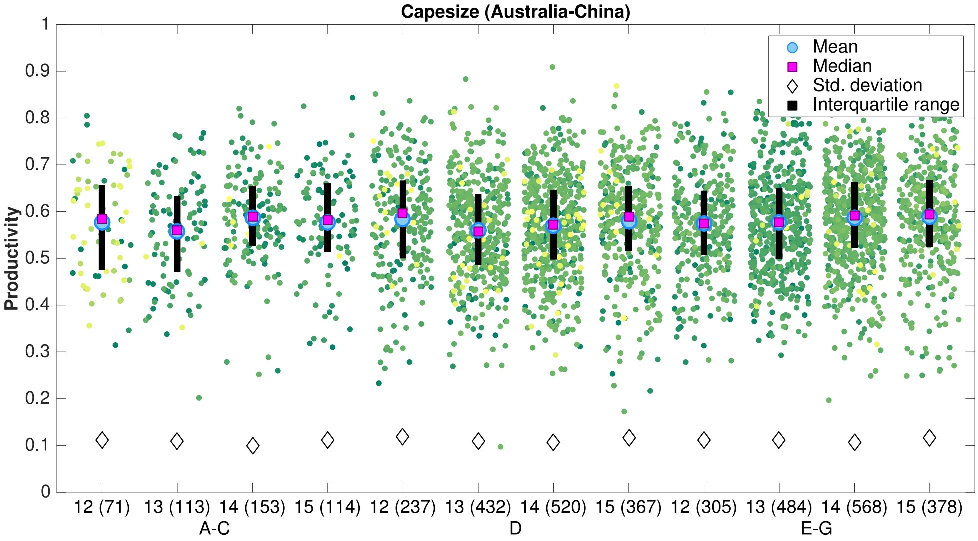

Figure 15: Productivity of the Capesize fleet on the Australia-China route

Pooling all years together, we find that median productivity for D-rated vessels is lower than the productivity of E-G rated vessels by 0.6%.

Median average ballast speed is lower by 0.4 knots in A-C compared to E-G, and 0.3 knots lower in D than E-G. Median average laden speed is slower by 0.2 knots in A-C than in E-G. There is also a trend of vessels slowing down over time, consistent with conditions in the market. The slow down is more pronounced at the voyage level than at the annual average level.

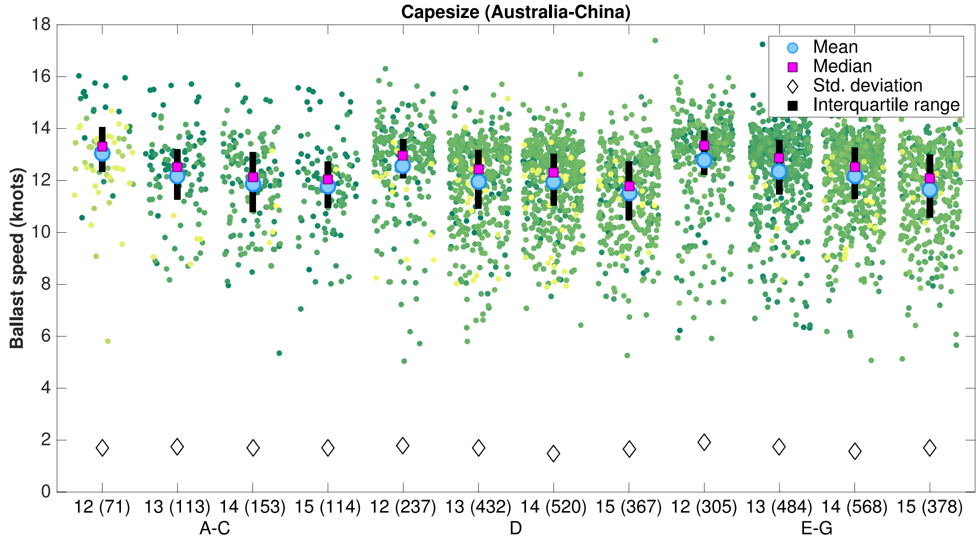

Figure 16: Ballast speeds of the Capesize fleet operating on the Australia-China route

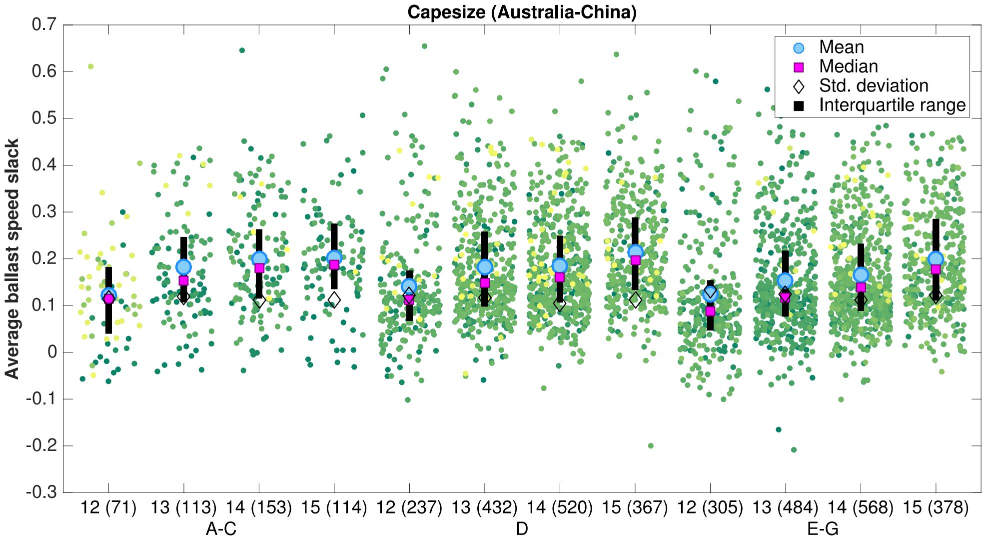

Figure 17: Ballast speed slack for the Capsize fleet operating between Australia and China

We get a similar set of results for voyages between Brazil and China, although the number of observations of this voyage is significantly smaller. Median ballast speed is 0.5 knots lower for A-C rated vessels compared to E-G, and about 0.3 knots slower in D than in E-G.

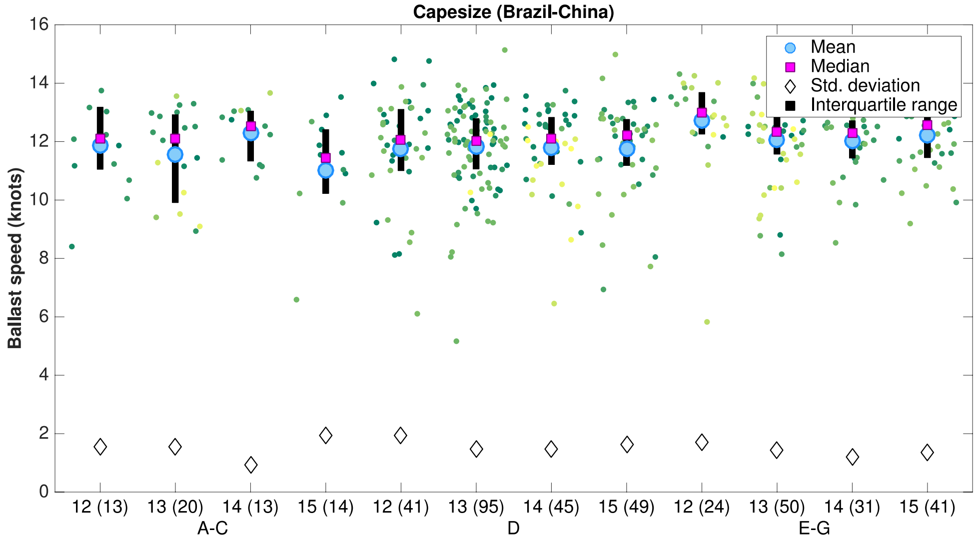

Figure 18: Ballast speed for the Capesize fleet operating between Brazil and China

Where we notice a much bigger difference is in productivity, with A-C rated vessels having 6% lower productivity (57%) compared to E-G (63%). A similar difference exists between A-C and D-rated vessels. Productivity is affected by sailing time (speed), distance, and capacity utilisation, although we only find significant differences in speed and a very small difference in capacity utilisation across the 3 GHG Rating categories. The small number of observations of this voyage might render the results less reliable.

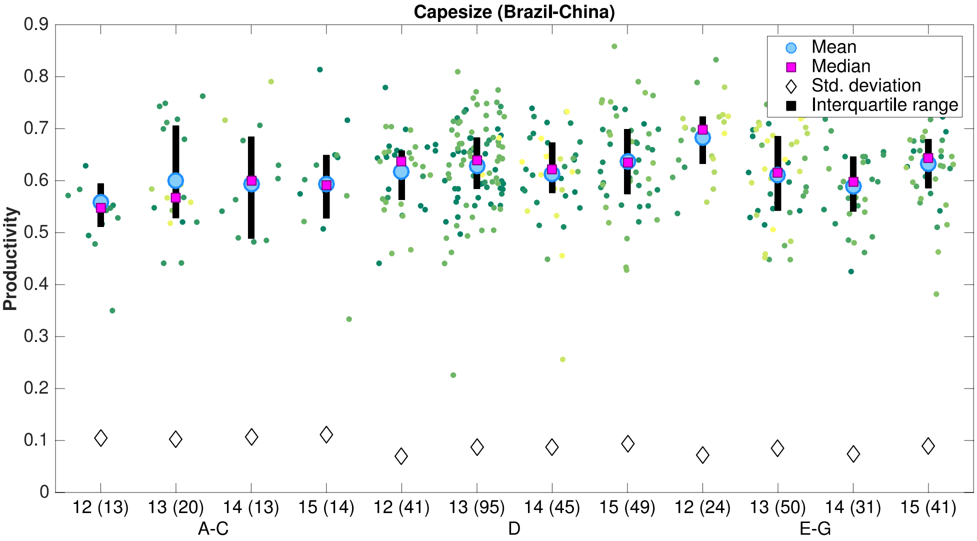

Figure 19: Productivity of the Capeszie fleet operating between Brazil and China

6.3.2 VLCC

Differences in productivity, ballast speeds, laden speeds, and the ballast time ratio are present in the in the top 10 voyages for VLCCs. However, the number of observations of voyages in each route is small and only account for a very small proportion of all VLCC voyages that must have taken place in each year between 2012 and 2015. If we assume a VLCC would complete a modest number of around 6 voyages a year (composed of a ballast leg followed by a laden leg), then a fleet of around

600 vessels ought to generate a sample of 3600 voyages a year. The current sample was never expected to capture a proportion close to the expected total per year, given the known issues in AIS data and the steps taken to verify and filter voyages, but the return here is far beyond what could be considered to be a sufficient or representative sample. Interpretation of the results here is therefore heavily cautioned.

Nonetheless, one trend spotted in the Capesize fleet with tonnage being operated more effectively on routes to Japan is present here, too. The ballast legs on voyages taking oil to Japan (from Saudi Arabia or UAE) are completed at speeds close to 14 knots at the median, whilst ballast journeys on other routes are significantly slower (between 10 and 12 knots).

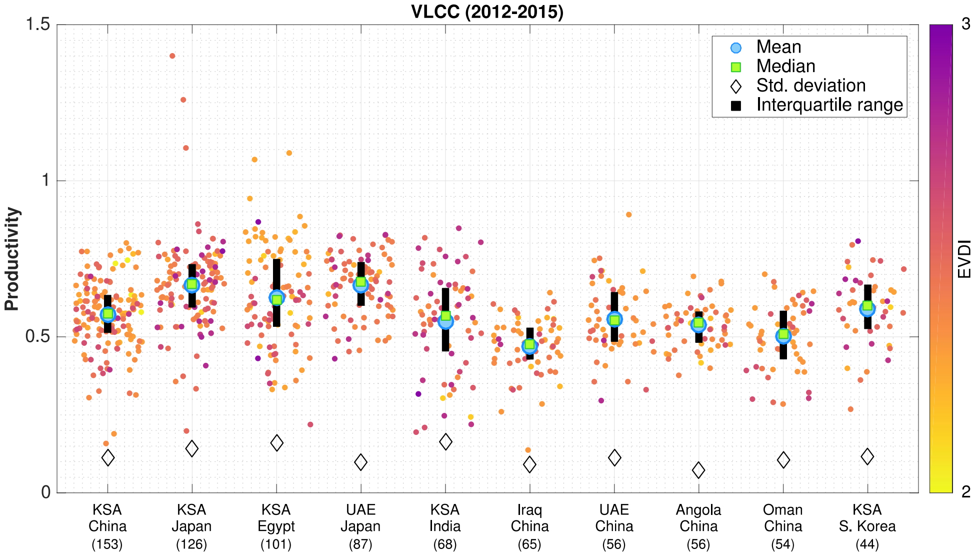

Figure 20: Productivity of the VLCC fleet by route

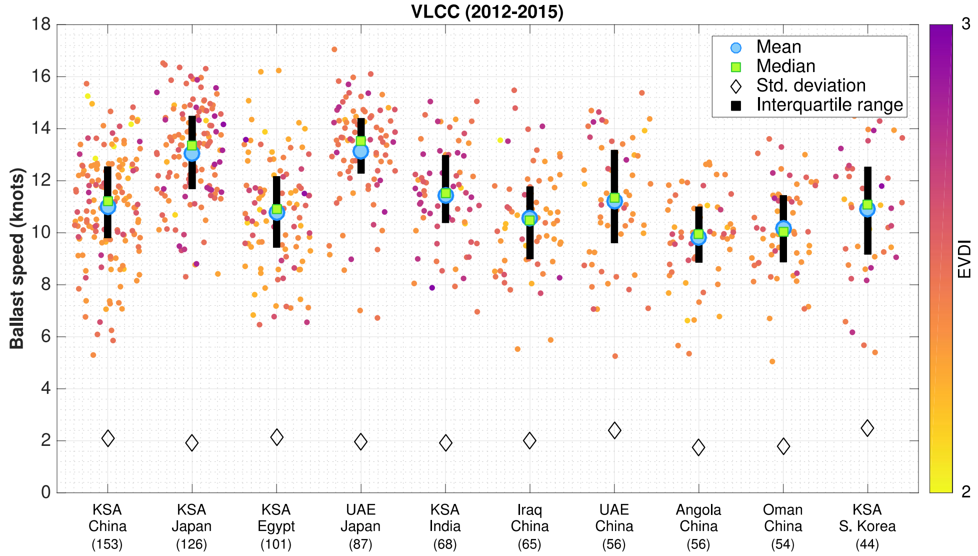

Figure 21: Ballast speeds of the VLCC fleet by route

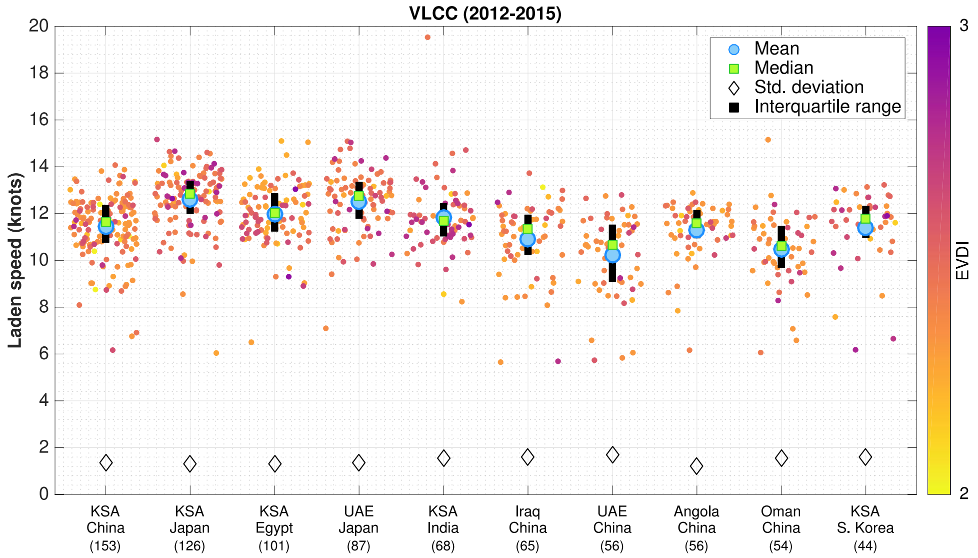

Figure 22: Laden speeds of the VLCC fleet by route

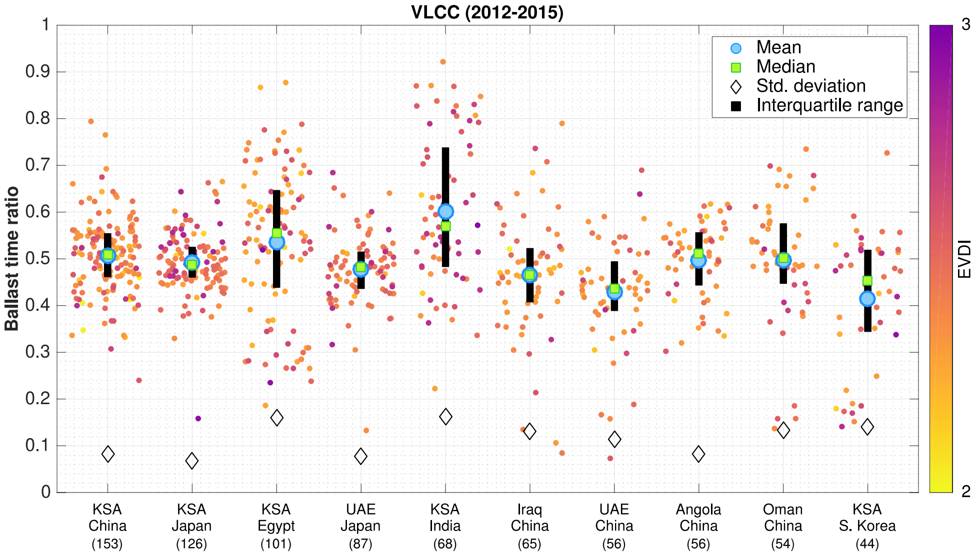

Figure 23: Ballast time ratios for the VLCC fleet by route

Unfortunately, the small sample size even when pooling all routes together imply that there are too few observations on even the most frequent route observed for VLCCs to perform a more detailed analysis. For example, splitting the voyages in the Saudi Arabia to China route by GHG Rating classes renders only on average about 6 voyages a year under the E-G category.

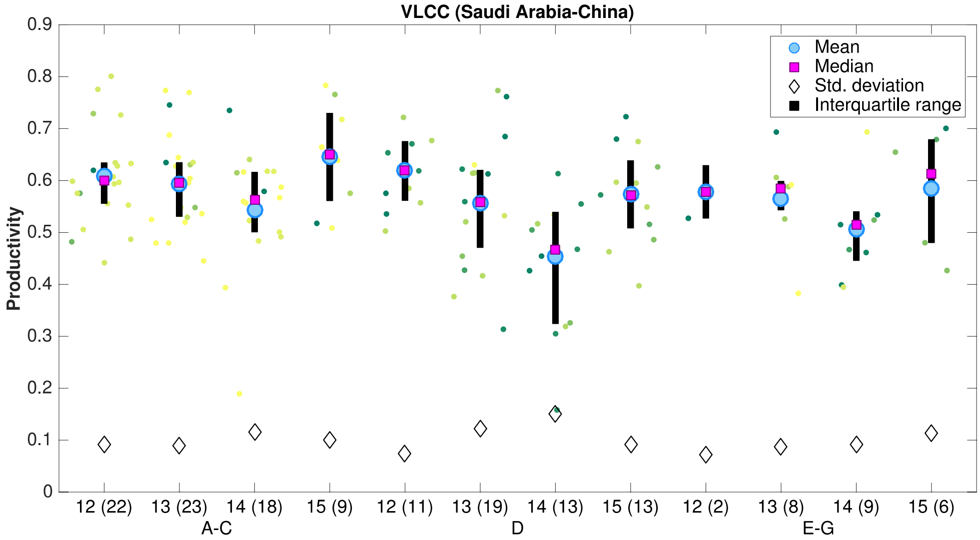

Figure 24: Productivity of the VLCC fleet by route

7 Summary

7.1 Price signals

We estimated whether vessels with good GHG Ratings were earning higher time charter rates than those with poor GHG Ratings using fixtures data from 2005 to 2015 for the Panamax (60,000–100,000 dwt) and Capesize (100,000+ dwt) dry bulk fleets and the Suezmax (120,000–200,000 dwt) and VLCC (200,000+ dwt) tanker fleets. The model, based on Parker & Prakash (2016) and Adland et al. (2015), accounted for vessel characteristics, local market conditions prior to and around the time of each fixture, the heterogeneity in time charter rates observed over time, space, and for different contract durations, and biases within the sample of fixtures available.

The model tested for the existence of a two-tier market for vessels based on their GHG Ratings, which was predicted to most likely to occur in the time charter market because the burden of fuel expenditure falls on the charterer rather than the ship owner. We looked to see if the difference in time charter rates was of a significant magnitude, as that could be considered a price signal or additional[23] incentive to invest in vessels with good GHG Ratings.

	Bulker		*Tanker*	
	Panamax	Capesize	Suezmax	VLCC
	(60–100k dwt)	*(100k+ dwt)*	*(120–200k dwt)*	*(200k+ dwt)*
Premium (%) received by ship owners in the time charter market for A-C rated vessels compared to D and above	2% ($1614/day max, $456/day mean, $70/day min)	0	0	0
Net savings in fuel cost per day to charterer in time charter (or owner in spot) from switching from an F to a B-rated vessel (*at design speed*)	$5823 (max), $3341 (mean), $580 (min)	$6130, $3922, $1556	$7333, $4555, $1717	$10843, $6736, $2539
% of observed fixtures in time charter (including time charter trips)	92.3	33.6	1.4	2.6
Total fixtures	12682	6303	12715	9076
Unique vessels	1773	923	446	601

Example vessels here are B and F-rated 75k dwt Panamaxes, 175k dwt Capesizes, 160k dwt Suezmaxes, and 310k dwt VLCCs.

Table 2: Summary of results (price signals)

7.1.1 Bulkers

1. For the Panamax fleet, we found that an A-C rated vessel returned on average a time charter day rate that was 2% higher than a D-rated vessel between 2005 and 2015. A-C rated vessels earned a similar premium over E-G rated vessels, such that there appeared not to be a sufficiently strong difference in the premium between D and E-G rated vessels.

---

[23] That is, in addition to the implicit saving in fuel costs.

2. There was no premium (or discount) for the Capesize fleet. This may be due to an insufficient sample of time charter fixtures across a sufficiently diverse group of vessels within the Capesize fleet—compared to the Panamax subset.
3. The most influential determinant of the time charter rate was the average of the time charter rates being issued close to but prior to a given fixture for similarly sized vessels, for similar contract durations, and for delivery in the same area, confirming the results derived in Adland et al. (2015). This method of including the average rate controlled for localised market conditions and rate differences across contract durations.
4. There was no indication that A-C rated vessels were more or less likely to be allocated to time charter contracts than spot contracts.

7.1.2 Tankers

1. A similar premium for A-C rated vessels was not observed with an acceptable level of statistical certainty for either the VLCC or Suexmax fleets.
2. Time charter fixtures represented less than 5% of all fixtures available between 2005 and 2015 in our dataset, which may have hindered our ability to detect a premium.
3. It remains unclear whether this allocation between spot and time charter is representative of the VLCC and Suezmax markets, since it is possible that our data source was biased towards the reporting of spot fixtures for these tanker sizes.
4. There is scope to test whether, under the right market conditions, tankers with good GHG Ratings are chartered in the spot market at a discount due to their implied lower fuel cost per mile.

7.2 Operational patterns

To establish whether vessels with good GHG Ratings were being operated differently to their peers with poor GHG Ratings, the operational performance of the Capesize and VLCC fleets from 2012 to 2015 was assessed at both an annual and voyage[24] level. We believed that differences in the way they were operated could have been an implicit sign that, even if a price signal was not observed, operators were potentially cognisant of its GHG Rating or the energy efficiency it proxies.

This was captured using AIS data, which provided information on a vessel's speed, position, draught, and a few other parameters at high frequency. The information was assimilated to classify vessels according to their loading condition (using draught), to estimate cargo tonnage, and identify voyages and routes (by detecting stops, establishing its location, chaining stops together, and connecting with loading conditions).

Several measures were defined to capture and summarise the operational performance of each vessel studied. At both an annual and a voyage level, these included:

1. Average ballast and laden speeds
2. 1 minus the ballast or laden speed as a ratio of the vessel's design speed (referred to as slack)
3. Time and distances over which the vessel was observed in ballast and laden
4. Capacity utilisation (the proportion of designed deadweight filled with cargo, where cargo is estimated using AIS-reported draught measurements)

---

[24] A voyage is defined as roughly a ballast leg followed by a laden leg, hence capturing the cost in time and distance required to reposition the vessel to load a new cargo.

A composite measure called productivity was created to compare voyages, defined as the ratio of the actual productivity (tonne nautical miles of work completed divided by the time it took for the whole voyage, including the ballast leg) to the potential or theoretical productivity (the amount of work or tonne nautical miles it could have completed had it sailed at full speed, carried maximum cargo, and had minimum idle time).

We statistically tested for significant differences in the median of each measure between A-C, D, and E-G rated vessels. Comparing medians was deemed a more robust—yet statistically powerful—approach in the presence of outliers and measurement errors inherited from AIS data and some of the estimation methods utilised (§6.1.1). Table 3 is a summary of the recorded median value of each measure for the A-C and E-G GHG Rating categories, along with an indication of whether the difference between any two medians was statistically significant.

	Capesize			VLCC		
Annual	A-C	E-G	Significant	A-C	E-G	Significant
Ballast speed	10.4	10.6	Y	9.7	10.3	Y
Laden speed	9.7	10	Y	10.2	11	Y
Ballast speed slack	0.29	0.27	Y	0.37	0.34	Y
Laden speed slack	0.33	0.31	Y	0.34	0.29	Y
Ballast time	0.374	0.373	N	0.411	0.408	N
Laden time	0.358	0.352	N	0.392	0.385	N
Ballast distance	0.51	0.51	N	0.5	0.5	N
Laden distance	0.49	0.49	N	0.5	0.5	N
Mean capacity utilisation	0.97	0.98	Y	0.92	0.91	Y

Table 3: Summary of results (annual operational patterns)

7.2.1 Capesize results

1. There is no indication that Capesize vessels in different rating categories were spending significantly different amounts of time in ballast or laden conditions annually.
2. Most of the significant differences occurred within measures of speed, but remained very small in magnitude. In general, there was little evidence to suggest GHG Ratings affected the way vessels were operated.
3. We did not make distinctions between the split in type of contracts any given vessel operated on annually or on specific routes, though contract type may be a determinant of operational performance.
4. For Capesize vessels, we found that median ballast speeds were lower by about 0.2 knots for A-C rated vessels and by 0.07 knots for D-rated vessels compared to E-G vessels over the 4 years analysed. Whilst the differences were statistically significant, they were notably small in magnitude. This effect where vessels with better ratings were sailing marginally slower extended to laden speeds, too.
5. Another way to look at differences in speed was to consider ballast or laden speed as a ratio of the vessel's design speed. Taking this ratio away from 1 gave us a measure of how far from design speed the vessel was actually being operated. We called this slack. Thus, a high % of slack in speed meant that the vessel was being operated further away from its design

speed. A-C rated vessels had 2% higher slack in annual average ballast speeds (29%) compared to E-G rated vessels (27%). That is, vessels with good GHG Ratings were being operated further away from their design speeds than those with poor GHG Ratings.

6. These tests were also run for specific years, avoiding the merging of different market conditions and allowing an exploration of the performance of the vessels in more comparable conditions. Of note, we find that slack in laden speed was 4% higher (32%) in A-C rated vessels than E-G (28%) in 2012, 2.5% higher in 2013, and 2.7% higher in 2014.
7. We also looked for differences in the measures between the 3 rating categories for two of the major Capesize routes. In particular, we evaluated voyages carrying cargo from Australia to China and from Brazil to China. As with the annual results, we found statistically significant differences in measures of speed for voyages between Australia and China, but the differences were small in magnitude. Typically, A-C rated vessels appeared to be operated slower than D or E-G rated vessels in both ballast and laden conditions.
8. There was extensive dispersion in most of the measures computed. That is, although medians and means appeared to be very similar across vessels in each of the 3 rating categories, the extent of variation within each category was generally high and there were a lot of observations beyond the interquartile range (and the upper and lower adjacent values). Some of this is due to measurement error, as the methods employed were contingent on the validity and reliability of data captured from AIS.

7.2.2 VLCC results

1. Relative to E-G rated vessels, we found that A-C rated vessels sail slower by about 0.5 knots in ballast and 0.7 knots slower whilst laden. This translated to a difference of 3% in ballast speed slack and a 6% difference in laden speed slack.
2. Speeds also differed between the 3 rating categories in specific years. For example, laden speeds for A-C rated vessels were slower by 1.2, 0.5, and 1.3 knots compared to E-G rated vessels in 2013, 2014, and 2015, respectively.
3. A small fleet, AIS draught values that were harder to validate and appeared to be infrequently updated, amongst other factors, made it difficult to derive as detailed results for the VLCC fleet as we were able to for the Capesize fleet.

7.3 Observations and explanations

Using all findings from this report, we look at where fuel savings accrue in the time charter market, estimate fuel savings for vessels typically operated on spot markets, suggest implications for stakeholders, and identify future work.

7.3.1 Price signal observations

We assess the implications of our results on the time charter rate difference between vessels with good and poor GHG Ratings by looking at the net saving in fuel costs for the charterer in the time charter market as well as the appropriation of total fuel cost savings between the charterer and ship owner. Net saving is defined as the fuel cost saving that accrues from using an A-C rated vessel compared to an E-G rated vessel less the premium paid in the time charter rate to obtain the better-rated vessel.

As shown in Table 1, our results indicate that, for example, switching from an F-rated Panamax bulk carrier to a B-rated alternative—equivalent in size, age, speed, and under similar market conditions—returns a peak net saving to the charterer of close to US$ 6000 per day if the vessel were contracted in markets where fuel prices were high and time charter rates were low. The mean net saving over 2005 to 2015 is close to US$ 3500 per day. Whilst the magnitude of the saving tends to vary, the net saving—net saving being fuel savings less the time charter premium—is always positive; so, it could be argued that, assuming everything else is equal, there is always an incentive for the charterer to hire vessels with good GHG Ratings. That is, the cost minimising solution for the charterer would always be to hire the vessel with the better GHG Rating[25].

The top plot in Figure 25 shows how the fuel savings from switching from this F-rated Panamax carrier to a B-rated equivalent varies over time. The middle plot shows the dollars per day of net fuel saving as a proportion of the time charter rate paid. As expected, daily fuel savings form a small proportion of around or less than 10% of the time charter day rate between 2005 and 2008, but increases up to 80% in poorer markets thereafter. The saving passed onto the ship owner peaks in markets where the time charter rates are high at about US$ 1600 per day in 2005 and 2008, equivalent to about 60% of the total fuel savings as shown in the bottom plot of Figure 25.

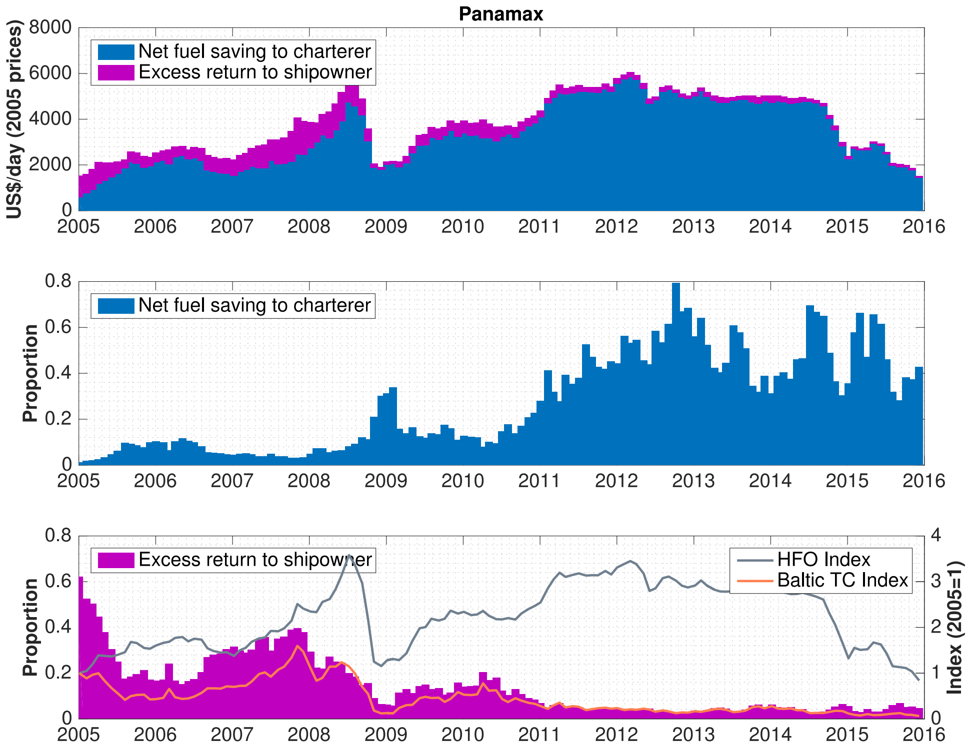

Figure 25: US$/day savings and excess return to ship owners (top), net savings to charterer as a proportion of the time charter rate (middle), excess return to ship owners as a proportion of total fuel savings (bottom)

Fuel savings exist also for the two tanker sizes considered. Although we were unable to establish a premium paid to ship owners in the time charter market for tankers, and given that tankers appear to be operated predominantly in the spot market, there are incentives to invest in vessels with good GHG Ratings as these could translate to substantial fuel cost savings to the operator.

---

[25] This depends on designed energy efficiency being an accurate proxy for operational energy efficiency, the availability of that information to parties involved, and the reliability of that information.

7.3.2 Operational pattern observations

Our results for the operational differences between vessels across our 3 GHG Rating categories can also be used to provide an expected difference in fuel costs, but not on the earnings of the vessels on an annual or voyage basis.

We typically found that, on average, Capesize (and VLCC) vessels in the A-C category tend to be operated more slowly than those in the D or E-G category. Consider a 175,000 deadweight-tonne Capesize vessel. In this vessel's peer group, a B, D, and F-rated vessel would have a gCO_2 per tonne nautical mile rating of 2.7, 2.9, and 3.1, respectively. At the A-C annual median ballast speed of 10.4 knots and a ballast draught of 9.5 metres, this translates to a tonnes per day fuel consumption rate of 13.8, 17, and 17.7, respectively[26]. If each vessel were instead operated at their measured median annual ballast speeds of 10.4, 10.5, and 10.6 knots, these consumption rates increase to 13.8, 17.5, and 18.7. Figure 26 below shows how this translates into fuel cost per day for these three Capesize vessels across the 4 years studied. The average HFO price per tonne in each year is included in parentheses in the *x*-axis.

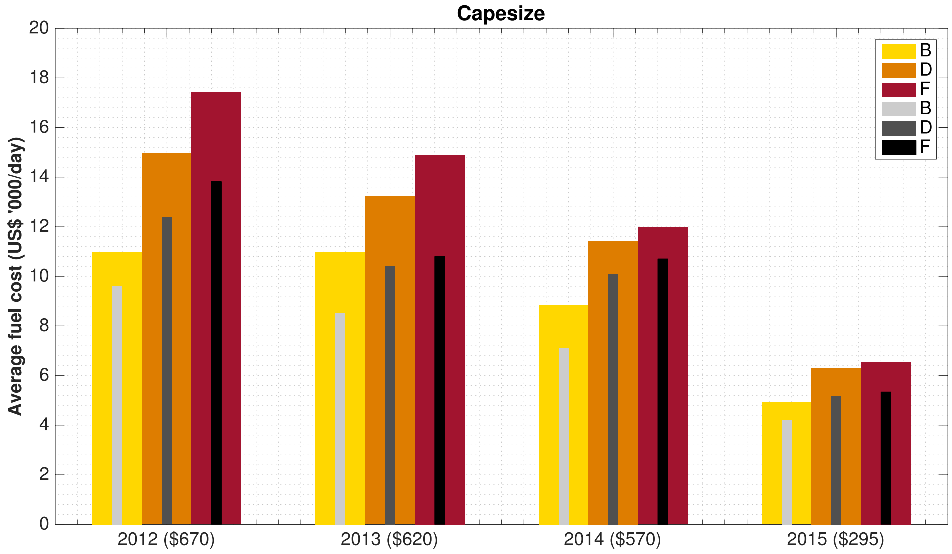

Figure 26: Fuel costs when laden (reds) and in ballast (greys) at annual average speeds

This case study suggests marginal differences in median annual speeds do not translate to significant differences in fuel consumption, even though there is an approximately cubic relationship between speed and fuel consumption (Third IMO GHG Study 2014), but instead the inherent fuel consumption differences between the B, D, and F-rated vessels produce the significant differences. However, we observed that vessels do travel faster (or slower) than the fleet-wide annual average ballast or laden speeds, so there are inevitably operators or charterers who would have benefitted more significantly from a vessel with a good GHG Rating than others. For example, annual average ballast speed varies between 11.5 and 9.5 knots in 2015 for both A-C and E-G rated vessels within the interquartile range alone (Figure 1).

---

[26] At their design speeds and ballast draughts, this equates to a TPD of 44.2, 46.1, and 50.9.

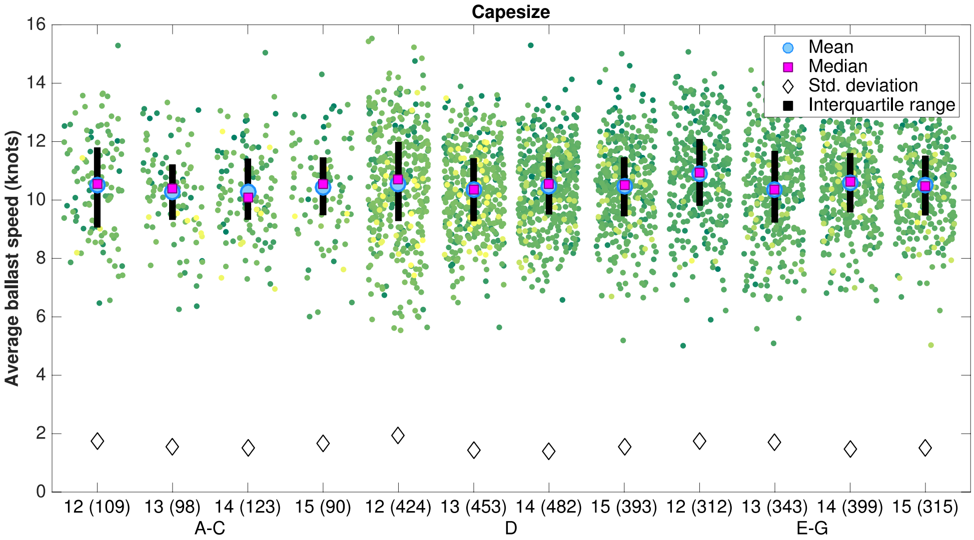

Figure 27: Average ballast speeds for the Capesize fleet

We had, however, expected that a B-rated vessel, for instance, would have been operated at a speed faster than a D or an F-rated vessel, given the delta in daily fuel consumption. This is because the B-rated Capesize vessel's annual average ballast speed could reach 11.25 knots before its daily fuel consumption equals that of a D-rated vessel operated at an annual average ballast speed of 10.5 knots. Our inability to detect this may purely be due to the markets being down in general between 2012 and 2015 at an *annual* level. There are, however, spikes in the market within each of the four years, such that our expected relationship may be observed if we worked at a higher resolution and were to model the speed of individual voyages as some function of GHG Ratings.

7.3.3 Candidate explanations for the pricing signal and operational patterns observed

We have identified several possible reasons as to why we have only been able to detect marginal differences in both the price signals and operational patterns of ships of varying GHG Ratings. Any one of these in isolation or some combination of these could explain the findings, but all provide suggestions for how future work could further help to understand the role of energy efficiency in shipping markets.

1. *Shortcomings in the datasets used for this analysis*

   Data quality in both fixtures and AIS tended to vary. Although we had a time period of 2005 to 2015 in fixtures data, we were not able to match all fixtures to vessels with a GHG Rating because GHG Ratings were only provided for vessels currently in service. This meant that there was a disproportionate loss in fixtures data for the early few years of the period covered.

   We also attempted to mitigate some of the biases inherent in the dataset of fixtures, because we knew that our data source did not have full coverage of all fixtures ever made or fixtures for every known vessel in the fleets studied. The true efficacy of the measures taken to correct biases can only be assessed if and when the model is applied to a new dataset wholly from another data source.

   AIS data and the measures derived using that data were also susceptible to errors, and AIS coverage increased non-linearly between 2012 and 2015. Although vessels with measures derived from AIS that appeared extreme were filtered from the sample before testing for differences between the 3 GHG Rating categories, many observations remained that could readily be classified as outliers. The test utilised is designed to mitigate potential problems from the presence of outliers.

2. *GHG Emissions Rating's validity*

   All of these results are conditional on the assumption that GHG Ratings and the gCO_2 per tonne nautical mile measure underlying them are:

   - Representative proxies for fuel efficiency or fuel consumption *in service and in reference* conditions
   - Such measures are likely to be used in the markets to assess chartering and pricing decisions
   - In the case of operational measures, the differences noted between the 3 rating categories are not representative of some other underlying structure (due to deadweight or cargo type, for example)
   - In the time charter model, the responsiveness of the time charter rate to changes in the rating category is not an artefact of some characteristic of each vessel that we have been unable to control for.

   Failure for any of these conditions could explain why the findings showed smaller differences between ships with different GHG Ratings to those that were expected.

3. *Limitations to using GHG Emissions Ratings*

   - The discrete nature of the GHG Ratings presented problems and limited the type of analysis possible, compared to a continuous measure like EVDI. However, EVDI is highly correlated with deadweight, and, the relative, peer-group specific nature of GHG Ratings allowed us to separate deadweight effects from relative efficiency effects.
   - Vessels were grouped into 3 rating categories, A to C, D, and E to G, which split the distribution of vessels well between good, average, and poor GHG Ratings (roughly 50% in D, and 25% in A-C and E-G). This had to be done to ensure that there were a sufficient number of fixtures for the price signal analysis and enough vessels for the AIS-derived operational analysis. Whilst this provided a statistically symmetric solution, there may be a discrepancy in the way vessels may be perceived in the market. For example, if only F and G-rated vessels are considered poor in terms of energy efficiency.
   - GHG Ratings provided were valid for the fleets studied in this report as of June 2016. This implies that vessels that were scrapped but still had fixtures for our price signal model from 2005 to 2015 or had AIS coverage for our operational measures from 2012 to 2015 were not included in the analysis.
   - Exclusion of those vessels does not appear to have biased our sample of vessels across the GHG Rating spectrum in either fixtures or AIS-derived data. That is, there is generally a consistent normal distribution of vessels across the 3 rating categories.
   - GHG Ratings are utilised under the assumption that they would have remained a valid proxy for relative energy efficiency regardless of when it was introduced or how prevalent it was. Thus, we rely on the assumption that vessels would have been compared in a similar manner regardless of whether this occurred before or after the introduction of the GHG Emissions Rating system.

4. *Small variations in EVDI*

   For certain fleets, we found that the typical difference in EVDI between two vessels in the same peer group between, say, an F and a B rating was very small. There was also not a lot of variation in EVDI within a given GHG Rating category. For example, the average EVDI values in the A-C, D, and E-G categories for the VLCC fleet were 2.4, 2.5, and 2.7, respectively, and the EVDI values within and across the 3 rating categories showed little variation (2–3% within each category, and 5% across all 3). However, the range of EVDI values in most fleets studied is reasonably wide.

5. *Timing, availability, and superseding conditions*

   At the point of charter, there are a limited number of vessels eligible for hire and it is unrealistic to assume that a charterer would wait for a vessel with a good GHG Rating to move into the area for charter. It may therefore not be feasible to always hire vessels with good GHG Ratings, and this inconsistency may have weakened or even nullified the signals we were trying to uncover. It may also be likely that conditions like the existence of sanctions, the vessel's safety record, and the role of brokers complicated or concealed the expected signals.

6. *Spot and time charter allocation*

   In our analysis of a vessel's operational patterns, we do not make distinctions about the contract types the vessel operated on over the year or on specific routes, even though contract type may be a driver of its operational performance (through the use of speed or fuel consumption clauses in the charter party, for example). It is possible that taking annual averages when vessels switch contract types masks or evens out differences in certain measures, although the contractual rigidity (particularly in the tanker markets) suggests this is rare.

   We did not cover or treat the spot market in the price signals part of this paper to the same level of detail. Extending our search for price signals that reveal a preference for vessels with good GHG Ratings to the spot markets is necessary to fully confirm the existence (or absence) of a premium or other signal for vessel types and sizes that predominantly operate on spot.

7.4 Conclusions

Both our time charter premium analysis and the assessment of operational behaviour showed that the advantage bestowed by a vessel with a good GHG Rating is mostly in the form of fuel savings. Hence, whom those savings accrue to and the magnitude and importance of those savings on his bottom line ought to characterise the incentive to invest in vessels with good GHG Ratings and low fuel consumption rates.

That is, these results suggest that investment decisions on vessels with good GHG Ratings are likely contingent on the expected allocation of the vessels between the spot and time charter markets, as well as expected fuel cost and freight rate conditions. Investments in vessels expected to spend most of their time in the time charter market may not be reliant on GHG Ratings if time charter rates are expected to be poor over the course of their lifespans or investment lifespan, since the excess returned to the ship owner is small and only prevalent in the Panamax dry bulk market. Whereas investment decisions on vessels that are expected to be operating predominantly on the spot market may be naturally more sensitive to GHG Ratings if fuel prices are expected to be high[27]. The size and nature of the influence of energy efficiency in spot markets remain avenues to be explored.

Assessing the additional investment cost involved in purchasing an A-C rated vessel rather than an E-G rated vessel (or the investment cost in converting an E-G to an A-C rated vessel) or other factors that affect the operating costs of the ship (maintenance costs, crewing, etc.) is necessary to provide a complete picture and could be a fruitful angle for future research, because it is possible that the potential discount in purchasing price and operating costs for an E-G rated vessel may outweigh the marginal excess returns to the ship owner from an A-C rated vessel.

Our ability to consistently estimate a premium in the time charter markets for the dry bulk and tanker fleets studied may have been hindered by the complexity and inherent inefficiencies of the markets themselves. For example, agents involved in contract negotiations may not have had reliable energy efficiency information, and such inherent market inefficiencies could have obscured the signals we were hoping to detect. Further, whilst the methods utilised to find a premium controlled for localised conditions in the market, vessel specifications, and the heterogeneity in the time charter rates observed across space, time, and contract length, there remain factors that either could not or have yet to be controlled for like a charterer's predisposition to a certain owner's vessels, the influence of brokers, as well as each vessel's safety records, sanctions, and proximity to the hire area.

In most of the operational measures computed for either the Capesize or VLCC fleets, there were no large differences in magnitude between the 3 rating categories at the median level and the implicit preference we had hoped to detect was not prevalent. There was, however, extensive variation within each category for most—if not all—operational measures. The existence of this variation is more interesting than the stability of the medians, as it suggests there were owners or charterers that either could have greatly benefitted from a vessel with a good GHG Rating or were already taking advantage of such a vessel by operating it more (or less) exhaustively. Identifying and modelling the drivers behind the variation on some of these measures remains an exciting opportunity for future work, and may reveal stronger signals than observed at the aggregate and average level taken in this report.

Underlying both sets of results is the GHG Emissions Rating system used as a proxy for relative energy efficiency, as is its applicability over a 10-year period. Whilst environmental concerns may not have been at the forefront of shipping markets 5 or 10 years ago, concerns over fuel consumption rates and costs ought to have always been present. The direct connection between GHG Ratings, EVDI, and fuel consumption should therefore be evident in the markets in some manner, and charterers ought to have been evaluating relative fuel costs for available vessels in a similar way to what is captured with GHG Ratings. That is, if GHG Emissions Ratings proxy energy efficiency, the existence of an explicit tool like the GHG Emissions Rating in older markets to help select vessels may not be critical to our results.

The ordinal nature of the GHG Ratings also presented a few modelling problems, but helped us better distinguish between deadweight and energy efficiency effects on time charter rates and operational performance measures. Vessels had to be grouped into 3 GHG Rating categories (A-C, D, E-G) to ensure that there were a sufficient number of fixtures and vessels from which to draw conclusions. GHG Ratings made available were also a snapshot of the active fleet as of June 2016, which meant that scrapped vessels were excluded and the vessel mix or peer groups could have

---

[27] There might, however, be cases where the revenue from very high freight rates renders fuel cost savings negligible or unimportant.

been different historically. Our tests on the consistency and shape of the distribution of vessels across the 3 GHG Rating categories over the 10 years studied suggest that there are no undue consequences to the results from using 2016 GHG Ratings retrospectively.

Despite these limitations, the results are encouraging for charterers. Charterers can readily exploit the weak or missing price signal where vessels with good GHG Ratings are not getting the implied higher efficiency recognised or accounted in the time charter day rates, by hiring vessels with good GHG Ratings and low fuel consumption rates without having to pay an added premium for the privilege. Whilst this may be advantageous in the short run, in the longer run this may become more costly. If the failure to reward ship owners for offering energy efficient vessels continues and the induction of increasingly more efficient vessels to the fleet falls as a result, charterers may be worse off in future markets where fuel prices are high and energy efficient vessels are in short supply.

However, the lack of strong preference signals on the basis of GHG Emissions Ratings also impacts how those ratings might be used by financiers looking to assess risks from potentially carbon-constrained future shipping markets. GHG Ratings do not provide an indication of exposure to other potential risks (technology obsolescence, for example) either. Together, this suggests that, in the short term, whilst there may be other benefits and the opportunity to influence wider behaviour, there may only be a modest portfolio-return incentive for financiers to make investment decisions using information on a ship's GHG Rating. Hence, a more sophisticated approach to assessing these risks may be necessary. For financiers, just as for owners, the need for high-quality, transparent information that can enable the shipping market to place a greater preference (premium or utilisation) on energy efficient ships remains an important priority.

Indeed, in spite of several years of availability and use of GHG Ratings, evidence of market failures and barriers remain. Regardless of expectations of future fuel prices or freight rates, this justifies the further use of regulation to assist in driving energy efficiency or decarbonisation. The effectiveness of price mechanisms to enable GHG reduction—through a carbon price or fuel levy, for example—may only be weak at incentivising the design of more efficient ships. Without first addressing the market barriers and failures implied in this report, extreme, artificial price signals may be required to achieve significant changes to vessel technology, operation, and GHG emissions.

The results generated in this report highlight the many issues present in both the way that shipping markets function when pricing GHG Emissions Ratings into transactions and the apparent discord between operational performance and GHG Emissions Ratings. The intricacies of the market as well as the presence of information barriers may have obfuscated the signals we had hoped to detect, such that addressing these fundamental issues and barriers within the markets may increase the importance of measures of as-designed relative energy efficiency. Nevertheless, a more comprehensive solution is necessary to properly address stranding risks, and we hope to get a better picture of what such a solution might look like when potential decarbonisation pathways are identified and explored in the next part of our stranded assets work.

8 References

Adland, R., Alger, H., and Banyte, J., 2015. Does fuel efficiency pay? Empirical evidence from the drybulk timecharter market revisited. *IAME 2015,* Kuala Lumpur.

Agnolucci, P., Smith, T. & Rehmatulla, N., 2014. Energy efficiency and time charter rates: Some evidence quantifying the extent of split incentive problem in the panamax drybulk market', *Transportation Research Part A: Policy and Practice*, *vol 66*, pp. 173–184.

Ansari, A. R. and Bradley, R. A., 1960. Rank-sum tests for dispersions. *The Annals of Mathematical Statistics*, *31*(4), pp.1174-1189.

Behrman, J. R., Wolfe, B. L., and Tunali, I. F., 1980. Determinants of women's earnings in a developing country: A double selectivity extended human capital approach. *Institute for Research on Poverty, University of Wisconsin, Madison.*

Catsiapis, G., and Robinson, C., 1982. Sample selection bias with multiple selection rules: an application to student aid grants. *Journal of Econometrics, 18 (3)*, pp. 351–368.

Faber, J., Behrends, B., Nelissen, D., 2011. Analysis of GHG Marginal Abatement Cost Curves. *CE Delft,* Delft.

Faber, J., Hoen, M., Koopman, M., Nelissen, D., Ahdour, S. 2015. Estimated Index Values of New Ships Analysis of EIVs of Ships That Have Entered The Fleet Since 2009. *CE Delft*, Delft.

Heckman, J. J., 1979, Sample selection bias as a specification error. *Econometrica, 47 (1),* pp. 153–162.

Jia, J., Prakash, V., and Smith, T., 2015. Estimating vessel utilisation in thedry bulk freight market. *UCL-NHH Working Papers*.

Kristensen, H. O., 2012. Determination of Regression Formulas for Main Dimensions of Tankers and Bulk Carriers based on IHS Fairplay data. *Project no. 2010-56, Emissionsbeslutningsstøttesystem, Work Package 2, Report no. 02, Technical University of Denmark*.

Kristensen, H. O., 2013. Statistical Analysis and Determination of Regression Formulas for Main Dimensions of Container Ships based on IHS Fairplay Data. *Project no. 2010- 56, Emissionsbeslutningsstøttesystem, Work Package 2, Report no. 03, Technical University of Denmark*.

Lloyds List, 2011, Owners urged to think long term on investment payback. Viewed 07 June 2011, *http://www.lloydslist.com/ll/sector/ship-operations/article372257.ece*.

Lloyds List, 2012, Playing with Fuel Performance. Viewed November 5 2012. *http://www.lloydslist.com/ll/sector/ship-operations/article410760.ece*.

Lloyds List, 2013, Who pays and who wins with fuel-saving retrofits? Viewed January 07 2013, *http://www.lloydslist.com/ll/sector/ship-operations/article414528.ece*.

Lloyds List, 2015, Poor performance tonnage rejected by 26 big charterers. *https://www.lloydslist.com/ll/sector/ship-operations/article514796.ece.*

Mitchell, J., and Rehmatulla, N., 2015. Dead in the water: an analysis of industry practices and perceptions on vessel efficiency and stranded ship assets. *2nd Shipping in Changing Climates conference*, Glasgow.

Parker, S. V., and Prakash, V., 2016. Energy efficiency premiums in the time charter market. *UCL Working Papers.*

Rehmatulla, N., 2014. Market failures and barriers affecting energy efficient operations in shipping (PhD thesis). *UCL Energy Institute, University College London (UCL),* London, United Kingdom.

Rightship, 2012. Cargill, Huntsman Corporation and Unipec UK set new industry standard on fuel efficiency for chartering vessels. URL: *http://site.rightship.com/resources/news-releases/2012/10/02/cargill,-huntsman-corporation-and-unipec-uk-set-new-industry-standard-on-fuel-efficiency-for-chartering-vessels/*

Rightship, 2015a. Calculating and Comparing CO_2 Emissions from the Global Maritime Fleet. URL: *http://www.shippingefficiency.org/sites/shippingefficiency.org/files/edvi-methodology.pdf*

Rightship, 2015b. COP21 - 2 billion tonnes of goods shipped using GHG Rating, URL: *http://site.rightship.com/resources/news-releases/2015/12/01/cop21-2-billion-tonnes-of-goods-shipped-using-ghg-rating/*

Smith, T., Traut, M., Bows-Larkin, A., Anderson, K., McGlade, C., Wrobel, P., 2015, CO_2 Targets, Trajectories and Trends for International Shipping.

Smith, T., O'Keeffe, E., Aldous, L., and Agnolucci, P., 2013. Assessment of Shipping's Efficiency Using Satellite AIS data. Prepared for the International Council on Clean Transportation. *UCL Energy Institute, University College London (UCL),* London, United Kingdom.

Smith, T., Prakash, V., Aldous, L., and Krammer, P., 2015, The existing fleet's CO_2 efficiency. IMO MEPC 68/INF.24/Rev.1.

Stopford, M., 2009. Maritime Economics, *3rd ed., Routledge*. New York, USA.

Wang, H., Faber, J., Nelissen, D., Russel, B., and St. Amand, D., 2010. Marginal Abatement Costs and Cost-Effectiveness of Energy Efficiency Measures. *Institute of Marine Engineering, Science & Technology (IMarEST)*. London, United Kingdom.

Wilcoxon, F., 1945 Individual comparisons by ranking methods. *Biometrics bulletin 1.6*, pp. 80–83.

Wilson, J., 2010. Carriage of Goods by Sea. *7th edition*, *Longman*. London, United Kingdom.

9 Tables

Table 4: Descriptive statistics of yearly summary distributions (Capesize)

	GHG Rating	Year	Vessels	Mean	Median	Std. dev.	Min	Max
Ballast time ratio	A-C	2012	109	0.41	0.40	0.13	0.01	0.76
		2013	98	0.38	0.35	0.10	0.04	0.67
		2014	123	0.37	0.37	0.09	0.16	0.77
		2015	90	0.40	0.39	0.09	0.21	0.67
	D	2012	424	0.40	0.38	0.11	0.07	0.74
		2013	453	0.35	0.35	0.09	0.00	0.63
		2014	482	0.36	0.36	0.08	0.00	0.67
		2015	393	0.40	0.39	0.10	0.11	0.75
	E-G	2012	312	0.41	0.40	0.11	0.05	0.77
		2013	343	0.35	0.35	0.09	0.02	0.68
		2014	399	0.36	0.36	0.08	0.01	0.69
		2015	315	0.41	0.41	0.10	0.06	0.68
Laden time ratio	A-C	2012	109	0.35	0.35	0.12	0.01	0.69
		2013	98	0.35	0.36	0.10	0.04	0.60
		2014	123	0.37	0.36	0.09	0.04	0.66
		2015	90	0.36	0.37	0.09	0.10	0.52
	D	2012	424	0.37	0.37	0.12	0.06	0.77
		2013	453	0.38	0.38	0.09	0.10	0.74
		2014	482	0.37	0.37	0.09	0.06	0.81
		2015	393	0.36	0.36	0.10	0.04	0.71
	E-G	2012	312	0.35	0.35	0.11	0.04	0.71
		2013	343	0.37	0.36	0.09	0.04	0.74
		2014	399	0.36	0.36	0.09	0.04	0.69
		2015	315	0.35	0.34	0.10	0.04	0.70

Ballast distance ratio	A-C	2012	109	0.54	0.52	0.16	0.01	0.99
		2013	98	0.52	0.52	0.12	0.08	0.94
		2014	123	0.50	0.50	0.11	0.20	0.94
		2015	90	0.52	0.52	0.11	0.24	0.81
	D	2012	424	0.52	0.51	0.13	0.06	0.90
		2013	453	0.50	0.50	0.11	0.00	0.79
		2014	482	0.50	0.50	0.11	0.00	0.88
		2015	393	0.53	0.52	0.12	0.15	0.94
	E-G	2012	312	0.54	0.53	0.12	0.08	0.92
		2013	343	0.50	0.51	0.12	0.03	0.87
		2014	399	0.50	0.50	0.11	0.01	0.93
		2015	315	0.54	0.53	0.11	0.04	0.89
Laden distance ratio	A-C	2012	109	0.46	0.48	0.16	0.01	0.99
		2013	98	0.48	0.48	0.12	0.06	0.92
		2014	123	0.50	0.50	0.11	0.06	0.80
		2015	90	0.48	0.48	0.11	0.19	0.76
	D	2012	424	0.48	0.49	0.13	0.10	0.94
		2013	453	0.50	0.50	0.11	0.21	1.00
		2014	482	0.50	0.50	0.11	0.12	1.00
		2015	393	0.47	0.48	0.12	0.06	0.85
	E-G	2012	312	0.46	0.47	0.12	0.08	0.92
		2013	343	0.50	0.49	0.12	0.13	0.97
		2014	399	0.50	0.50	0.11	0.07	0.99
		2015	315	0.46	0.47	0.11	0.11	0.96
Average ballast speed	A-C	2012	109	10.53	10.54	1.73	6.47	15.29
		2013	98	10.26	10.39	1.56	6.26	13.34
		2014	123	10.29	10.09	1.50	6.95	15.05
		2015	90	10.44	10.53	1.67	6.01	14.30

	D	2012	424	10.57	10.70	1.93	5.54	15.53
		2013	453	10.37	10.34	1.45	5.64	13.89
		2014	482	10.48	10.56	1.38	6.58	15.30
		2015	393	10.46	10.52	1.54	5.19	15.01
	E-G	2012	312	10.90	10.95	1.73	5.01	15.07
		2013	343	10.34	10.34	1.71	5.09	14.45
		2014	399	10.59	10.63	1.47	6.56	14.18
		2015	315	10.52	10.49	1.50	5.03	14.29
Average laden speed	A-C	2012	109	9.65	9.75	1.85	5.71	14.01
		2013	98	9.79	10.00	1.27	6.64	12.72
		2014	123	9.53	9.60	1.40	5.58	12.77
		2015	90	9.76	9.81	1.21	7.18	12.38
	D	2012	424	9.94	10.08	1.63	5.23	14.43
		2013	453	9.78	9.91	1.31	5.14	13.70
		2014	482	9.67	9.75	1.23	5.15	12.89
		2015	393	9.86	9.94	1.25	5.20	13.47
	E-G	2012	312	10.13	10.47	1.68	5.03	13.69
		2013	343	10.00	10.18	1.42	5.56	13.43
		2014	399	9.81	9.77	1.20	6.34	13.01
		2015	315	9.84	9.93	1.24	5.27	12.77
Average capacity utilisation	A-C	2012	109	0.95	0.96	0.07	0.63	1.30
		2013	98	0.96	0.97	0.06	0.69	1.11
		2014	123	0.97	0.97	0.05	0.78	1.10
		2015	90	0.96	0.98	0.05	0.78	1.11
	D	2012	424	0.96	0.97	0.06	0.46	1.19
		2013	453	0.96	0.97	0.06	0.69	1.18
		2014	482	0.96	0.97	0.06	0.69	1.19
		201	393	0.97	0.97	0.06	0.8	1.49

		5					1	
	E-G	2012	312	0.97	0.98	0.06	0.47	1.17
		2013	343	0.97	0.98	0.07	0.62	1.17
		2014	399	0.97	0.98	0.06	0.62	1.48
		2015	315	0.97	0.98	0.06	0.75	1.21
Ballast speed slack	A-C	2012	109	0.28	0.27	0.12	-0.01	0.57
		2013	98	0.31	0.30	0.11	0.09	0.59
		2014	123	0.30	0.30	0.10	0.00	0.54
		2015	90	0.29	0.28	0.12	0.05	0.64
	D	2012	424	0.28	0.27	0.13	-0.06	0.64
		2013	453	0.29	0.30	0.10	0.01	0.63
		2014	482	0.28	0.28	0.10	-0.04	0.55
		2015	393	0.28	0.28	0.11	-0.01	0.64
	E-G	2012	312	0.25	0.25	0.12	0.00	0.64
		2013	343	0.29	0.29	0.12	-0.03	0.64
		2014	399	0.27	0.27	0.10	-0.04	0.55
		2015	315	0.28	0.28	0.11	-0.04	0.67
Laden speed slack	A-C	2012	109	0.35	0.32	0.13	0.05	0.63
		2013	98	0.34	0.33	0.09	0.11	0.56
		2014	123	0.35	0.35	0.10	0.15	0.63
		2015	90	0.33	0.34	0.09	0.09	0.53
	D	2012	424	0.32	0.31	0.11	0.04	0.66
		2013	453	0.33	0.33	0.09	0.08	0.65
		2014	482	0.34	0.33	0.09	0.06	0.65
		201	393	0.33	0.32	0.09	0.0	0.66

		5					5	
	E-G	2012	312	0.30	0.28	0.12	-0.01	0.65
		2013	343	0.31	0.30	0.10	0.10	0.63
		2014	399	0.33	0.32	0.09	0.07	0.56
		2015	315	0.32	0.32	0.09	0.06	0.64

Table 5: Descriptive statistics of yearly summary distributions (VLCC)

	GHG Rating	Year	Vessels	Mean	Median	Std. dev.	Min	Max
Ballast time ratio	A-C	2012	65	0.44	0.45	0.12	0.10	0.79
		2013	74	0.41	0.41	0.13	0.01	0.66
		2014	94	0.40	0.41	0.10	0.18	0.68
		2015	49	0.36	0.36	0.14	0.08	0.64
	D	2012	68	0.41	0.42	0.14	0.09	0.76
		2013	104	0.42	0.43	0.13	0.05	0.74
		2014	133	0.40	0.41	0.12	0.00	0.69
		2015	62	0.37	0.36	0.14	0.11	0.64
	E-G	2012	61	0.42	0.42	0.13	0.09	0.74
		2013	65	0.41	0.43	0.13	0.12	0.76
		2014	85	0.39	0.39	0.10	0.14	0.62
		2015	39	0.40	0.42	0.12	0.15	0.61
Laden time ratio	A-C	2012	65	0.40	0.39	0.13	0.07	0.79
		2013	74	0.40	0.38	0.13	0.11	0.79
		2014	94	0.38	0.38	0.11	0.16	0.63
		2015	49	0.44	0.42	0.15	0.16	0.72
	D	2012	68	0.42	0.42	0.15	0.15	0.82
		2013	104	0.40	0.40	0.12	0.14	0.78
		2014	133	0.39	0.39	0.14	0.07	0.87
		2015	62	0.41	0.39	0.14	0.14	0.75
	E-G	2012	61	0.39	0.39	0.15	0.10	0.79
		2013	65	0.40	0.38	0.13	0.12	0.77
		2014	85	0.39	0.38	0.11	0.07	0.65
		2015	39	0.40	0.40	0.13	0.14	0.68
Ballast distance ratio	A-C	2012	65	0.51	0.51	0.14	0.11	0.85
		2013	74	0.48	0.50	0.15	0.01	0.75
		2014	94	0.48	0.50	0.13	0.13	0.79
		2015	49	0.44	0.45	0.16	0.03	0.86
	D	2012	68	0.47	0.49	0.18	0.12	0.84
		2013	104	0.50	0.49	0.15	0.05	0.85
		2014	133	0.49	0.51	0.15	0.00	1.00
		2015	62	0.47	0.50	0.16	0.05	0.76

	E-G	2012	61	0.50	0.49	0.18	0.09	0.87
		2013	65	0.48	0.52	0.15	0.13	0.84
		2014	85	0.48	0.49	0.13	0.09	0.75
		2015	39	0.48	0.51	0.17	0.18	0.81
Laden distance ratio	A-C	2012	65	0.49	0.49	0.14	0.15	0.89
		2013	74	0.52	0.50	0.15	0.25	0.99
		2014	94	0.52	0.50	0.13	0.21	0.87
		2015	49	0.56	0.55	0.16	0.14	0.97
	D	2012	68	0.53	0.51	0.18	0.16	0.88
		2013	104	0.50	0.51	0.15	0.15	0.95
		2014	133	0.51	0.49	0.15	0.00	1.00
		2015	62	0.53	0.50	0.16	0.24	0.95
	E-G	2012	61	0.50	0.51	0.18	0.13	0.91
		2013	65	0.52	0.48	0.15	0.16	0.87
		2014	85	0.52	0.51	0.13	0.25	0.91
		2015	39	0.52	0.49	0.17	0.19	0.82
Average ballast speed	A-C	2012	65	10.35	10.54	1.83	5.65	13.77
		2013	74	9.37	9.36	1.69	6.24	13.49
		2014	94	9.62	9.58	1.41	6.23	13.86
		2015	49	10.15	10.22	1.41	5.92	12.90
	D	2012	68	10.24	10.06	2.01	6.14	14.38
		2013	104	9.84	9.69	1.94	5.58	14.84
		2014	133	10.09	10.08	1.73	5.04	13.85
		2015	62	10.39	10.30	1.74	7.13	15.43
	E-G	2012	61	10.14	10.15	1.93	5.79	14.97
		2013	65	9.71	9.51	2.03	5.87	14.03
		2014	85	10.50	10.41	1.55	7.49	13.72
		2015	39	10.43	10.42	1.50	6.61	12.90
Average laden speed	A-C	2012	65	10.52	10.88	1.94	5.98	13.79
		2013	74	10.09	10.07	1.69	5.43	13.35
		2014	94	10.07	10.11	1.58	5.22	12.95
		2015	49	10.09	10.06	1.52	6.58	12.79
	D	2012	68	10.91	11.04	1.87	5.48	13.78
		2013	104	10.63	10.81	1.68	6.53	13.84
		2014	133	10.72	10.73	1.45	6.80	13.82
		2015	62	10.58	10.76	1.66	6.32	13.51
	E-G	2012	61	10.96	11.24	1.73	6.73	13.85
		2013	65	10.66	11.28	1.81	5.65	13.73
		2014	85	10.62	10.58	1.29	6.69	13.72
		2015	39	10.98	11.40	1.30	8.33	12.91
Average capacity utilisation	A-C	2012	65	0.92	0.92	0.06	0.77	1.08
		2013	74	0.91	0.92	0.08	0.52	1.07
		2014	94	0.93	0.93	0.06	0.80	1.10
		2015	49	0.91	0.93	0.08	0.61	1.07
	D	2012	68	0.91	0.92	0.05	0.76	1.01

		2013	104	0.89	0.91	0.06	0.61	1.00
		2014	133	0.89	0.90	0.07	0.54	1.02
		2015	62	0.91	0.92	0.06	0.70	1.00
	E-G	2012	61	0.90	0.90	0.05	0.75	0.99
		2013	65	0.91	0.91	0.06	0.78	1.05
		2014	85	0.91	0.91	0.07	0.74	1.07
		2015	39	0.90	0.91	0.08	0.57	1.04
Ballast speed slack	A-C	2012	65	0.34	0.33	0.12	0.10	0.65
		2013	74	0.40	0.40	0.11	0.10	0.60
		2014	94	0.38	0.39	0.09	0.05	0.59
		2015	49	0.35	0.33	0.09	0.17	0.63
	D	2012	68	0.35	0.35	0.13	0.06	0.60
		2013	104	0.37	0.38	0.13	0.03	0.65
		2014	133	0.36	0.36	0.11	0.13	0.67
		2015	62	0.34	0.34	0.11	0.05	0.54
	E-G	2012	61	0.34	0.35	0.13	0.03	0.62
		2013	65	0.37	0.38	0.13	0.11	0.63
		2014	85	0.32	0.32	0.10	0.11	0.55
		2015	39	0.33	0.33	0.09	0.16	0.56
Laden speed slack	A-C	2012	65	0.33	0.31	0.12	0.10	0.61
		2013	74	0.35	0.35	0.11	0.09	0.66
		2014	94	0.35	0.34	0.11	0.11	0.67
		2015	49	0.35	0.35	0.10	0.18	0.57
	D	2012	68	0.30	0.29	0.12	0.13	0.65
		2013	104	0.32	0.31	0.11	0.10	0.59
		2014	133	0.32	0.31	0.09	0.11	0.56
		2015	62	0.32	0.30	0.10	0.13	0.59
	E-G	2012	61	0.29	0.27	0.11	0.09	0.56
		2013	65	0.31	0.27	0.12	0.13	0.61
		2014	85	0.31	0.31	0.08	0.13	0.53
		2015	39	0.29	0.28	0.08	0.15	0.47

Table 6: Wilcoxon rank-sum results for differences in medians between EVDI classes (Capesize)

			Median			*Vessels*		
	(1)	*(2)*	*(1)*	*(2)*	*%Δ*	*(1)*	*(2)*	*P-value*
Average ballast speed	A-C	D	10.39	10.52	1.2	420	1752	0.27
	A-C	E-G	10.39	10.60	2	420	1369	0.02
	D	E-G	10.52	10.60	0.72	1752	1369	0.04
Average ballast speed slack	A-C	D	0.29	0.28	-2.8	420	1752	0.20
	A-C	E-G	0.29	0.27	-6.9	420	1369	0.00
	D	E-G	0.28	0.27	-4.2	1752	1369	0.00
Average capacity utilisation	A-C	D	0.97	0.97	0.15	420	1752	0.33
	A-C	E-G	0.97	0.98	1	420	1369	0.00
	D	E-G	0.97	0.98	0.85	1752	1369	0.00

Average laden speed	A-C	D	9.78	9.88	1	420	1752	0.12
	A-C	E-G	9.78	10.02	2.5	420	1369	0.00
	D	E-G	9.88	10.02	1.4	1752	1369	0.00
Average laden speed slack	A-C	D	0.33	0.33	-2.6	420	1752	0.04
	A-C	E-G	0.33	0.31	-6.9	420	1369	0.00
	D	E-G	0.33	0.31	-4.3	1752	1369	0.00
Ballast distance ratio	A-C	D	0.51	0.51	-0.75	420	1752	0.72
	A-C	E-G	0.51	0.51	0.055	420	1369	0.34
	D	E-G	0.51	0.51	0.82	1752	1369	0.05
Ballast time ratio	A-C	D	0.37	0.37	-1.3	420	1752	0.18
	A-C	E-G	0.37	0.37	-0.32	420	1369	0.50
	D	E-G	0.37	0.37	0.98	1752	1369	0.33
Laden distance ratio	A-C	D	0.49	0.49	0.79	420	1752	0.72
	A-C	E-G	0.49	0.49	-0.058	420	1369	0.34
	D	E-G	0.49	0.49	-0.85	1752	1369	0.05
Laden time ratio	A-C	D	0.36	0.37	2.7	420	1752	0.03
	A-C	E-G	0.36	0.35	-1.8	420	1369	0.54
	D	E-G	0.37	0.35	-4.4	1752	1369	0.00

The null hypothesis that the medians are equivalent is rejected if the P-value is below 0.05.

Table 7: Wilcoxon rank-sum results for differences in medians between EVDI classes (VLCC)

			Median			*Vessels*		
	(1)	*(2)*	(1)	(2)	*%Δ*	(1)	(2)	*P-value*
Average ballast speed	A-C	D	9.72	10.06	3.6	282	367	0.05
	A-C	E-G	9.72	10.27	5.7	282	250	0.01
	D	E-G	10.06	10.27	2.1	367	250	0.41
Average ballast speed slack	A-C	D	0.38	0.36	-5	282	367	0.13
	A-C	E-G	0.38	0.34	-9.8	282	250	0.00
	D	E-G	0.36	0.34	-5.1	367	250	0.08
Average capacity utilisation	A-C	D	0.92	0.91	-1.7	282	367	0.00
	A-C	E-G	0.92	0.91	-2	282	250	0.00
	D	E-G	0.91	0.91	-0.32	367	250	0.76
Average laden speed	A-C	D	10.17	10.78	6	282	367	0.00
	A-C	E-G	10.17	11.01	8.2	282	250	0.00
	D	E-G	10.78	11.01	2.1	367	250	0.57
Average laden speed slack	A-C	D	0.34	0.31	-9.5	282	367	0.00
	A-C	E-G	0.34	0.29	-16	282	250	0.00
	D	E-G	0.31	0.29	-7.6	367	250	0.09
Ballast distance ratio	A-C	D	0.50	0.50	-0.61	282	367	0.77
	A-C	E-G	0.50	0.50	0.65	282	250	0.76
	D	E-G	0.50	0.50	1.3	367	250	0.96
Ballast time ratio	A-C	D	0.41	0.41	-0.91	282	367	0.84
	A-C	E-G	0.41	0.41	-0.69	282	250	0.97
	D	E-G	0.41	0.41	0.22	367	250	0.98

Laden distance ratio	A-C	D	0.50	0.50	0.61	282	367	0.77
	A-C	E-G	0.50	0.50	-0.65	282	250	0.76
	D	E-G	0.50	0.50	-1.3	367	250	0.96
Laden time ratio	A-C	D	0.39	0.40	2	282	367	0.87
	A-C	E-G	0.39	0.39	-1.9	282	250	0.85
	D	E-G	0.40	0.39	-3.8	367	250	0.73

The null hypothesis that the medians are equivalent is rejected if the P-value is below 0.05.