# The Information Constraint in Market-based Sanctions Enforcement*

Avi Dutt, Abhiroop Mukherjee, George Panayotov, Debjit Roy, Xudong Wen

May 23, 2026

**Abstract**

Why do sanctions violations remain pervasive, even as authorities now enlist third-party market intermediaries to exclude violators from global finance and trade? We show that the primary reason is not weak compliance incentives — as commonly assumed in policy debates — but rather the difficulty these third parties face in identifying violators. Our analysis exploits the sudden disclosure of a list of suspected sanctions-violating oil tankers on Refinitiv Eikon, the world's second-most widely used business data platform, by the maritime analytics firm Windward.ai. This disclosure increased third-party detection accuracy (pseudo-$R^2$) from 18% to 42%, resulting in a 13% decline in earnings for suspect tankers, a 17% drop in their likelihood of approaching sanctioned countries, and a 37% drop in their resale probability. A dynamic structural model calibrated to these market impacts implies that the information shock redirected $1.1 billion a year away from violators and reduced compliant exporters' shipping costs by 3.5%, but it also reduced compliant tankers' earnings. The model further explains why Windward – despite the sizable market impact of its list – was valued at only $42 million just before the disclosure. Finally, counterfactual analysis confirms that stronger penalties alone are ineffective without better information.

**Keywords**: Sanctions violations, Information disclosure, Oil shipping, Market failure

**JEL codes**: F51, D82, G14

*Dutt: Trademo; Mukherjee, Panayotov, and Wen: Hong Kong University of Science and Technology; Roy: Indian Institute of Management Ahmedabad (IIMA). We are particularly grateful to Nick Barberis, Briana Chang, and Amit Seru for their encouragement and comments, and thank Sumit Agarwal, Tania Babina, Utpal Bhattacharya, Darwin Choi, Shan Ge, Pulak Ghosh, John Griffin, Minkang Gu, Harrison Hong, Yunzhi Hu, Yan Ji, Navin Kumar, Erica Li, Vikram Mathur, Jagmeet Makkar, S. Lakshmi Naaraayanan, Kasper Nielsen, John Nash, Jun Pan, Elias Papaioannou, Krishna Prasad, Robert Ready, Vivek Reshamwala, Arkodipta Sarkar, Rik Sen, Janghoon Shon, Ashok Srinivasan, Yongxiang Wang, Constantine Yannelis, Jianfeng Yu, Xiaoyun Yu, Alminas Zaldokas, Jian Zhang and seminar participants and students at CKGSB, European Economic Association Conference 2025, EduHK, FMA Asia, HKU Summer Finance Conference, HKUST, London Business School, Monash University, SAIF, and Tsinghua University for various helpful comments and suggestions. We also thank the UN Global Platform for access to its AIS database (UN Statistics Division, "Global AIS Data Feed.", 2020), Vishesh Vatsal and Dfy-graviti for their help in analyzing satellite imagery, and Niranjana Unnithan and Benas Zurauskas for excellent research assistance.

In recent decades, sanctions enforcement has evolved beyond an exclusive reliance on state-imposed penalties toward a more decentralized, market-based framework. In this system, sanctioning authorities require third-party intermediaries — such as banks, insurers, shipping companies, and other service providers — to conduct due diligence, monitor counterparties, and terminate relationships with entities that facilitate violations (see, e.g., Hufbauer and Jung, 2020; Norrlöf, 2021; Van Genugten, 2021). The scope of sanctions has thus been broadened by enlisting non-state actors to restrict violators' access to critical financial and logistical infrastructure — such as the SWIFT payment network or the U.S. dollar-based financial system. Yet enforcement outcomes remain uneven, with continued and widespread violations revealing significant challenges (see, e.g., Fisman et al. (2024); ESRC's *Economic Observatory*, or, *The New Yorker*).

In this paper, we study whether these enforcement challenges arise primarily from information constraints – i.e., because third parties have difficulty identifying violators — or from limited compliance incentives — i.e., they can identify violators but nonetheless continue dealing with them.

Sanctioning authorities like the U.S. Office of Foreign Assets Control (OFAC) have typically focused on strengthening incentives, by imposing steep penalties for non-compliance (e.g., BNP Paribas paid a $8.97 bln fine in 2014), extending statutes of limitations from five to ten years (OFAC, July 2024), and clarifying that sanctions operate on a strict liability basis, meaning enforcement of penalties does not require proof of intent (OFAC, March 2024).[1]

However, sanctioning authorities provide only limited support in identifying violators, effectively shifting the burden of detection onto market participants. Consider the oil shipping sector, which plays a central role in global trade (Brancaccio, Kalouptsidi and Papageorgiou, 2020, 2023; Ready, 2018) and serves as the primary conduit for vital oil export revenues of sanctioned states such as Russia and Iran (Brown, 2020). In this sector, U.S. sanctioning authorities added approximately 100 tankers to publicly available sanctioned entity lists between 2021 and 2024. Over the same period, however, the Congressional Research Service estimated that more than 1,600 vessels transported sanctioned oil, implying that identification of the remaining 1,500 tankers was left to third parties.

At the same time, market participants across sectors increasingly underscore the difficulty of detecting sanctions violators. In finance, for instance, a review of 10-K filings by the 50 largest U.S. institutions shows that 97% explicitly cite the challenge of identifying violators as a significant compliance risk (Figure 1).

[1] This focus on incentives is not unique to the U.S.: other sanctioning authorities such as the EU and the UK have adopted comparable enforcement principles; see, e.g., FCA, 2024).

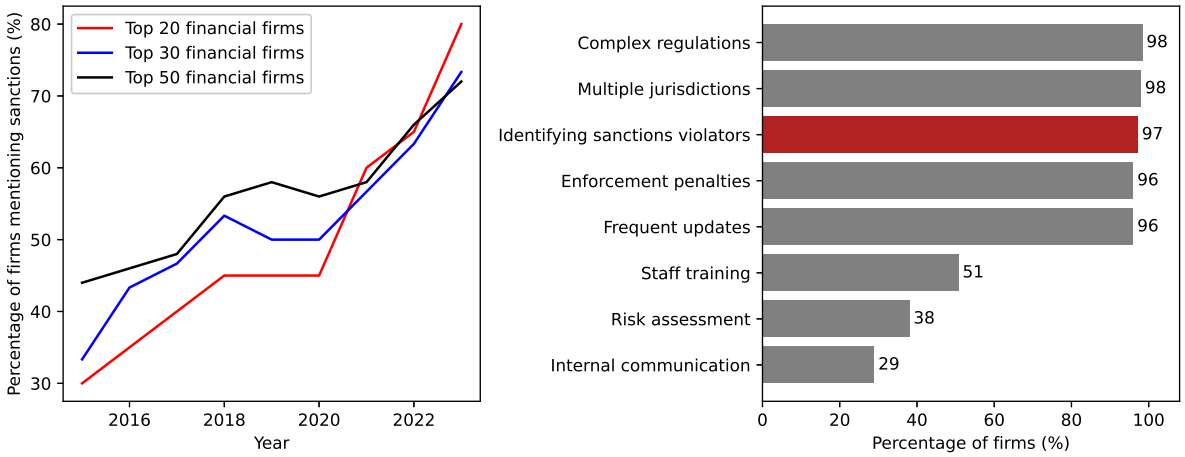



Figure 1: This figure shows the results from a large-language-model-based (OpenAI's ChatGPT-4) analysis of 10-K reports of the largest (by total assets) U.S. financial firms between 2015 and 2023. We first ask GPT-4 to identify the reports that mention sanctions enforced by OFAC, or sanctions-related risk, cost, or uncertainty. Then, we ask it to identify the reasons why the firms think sanctions compliance is challenging (further details are in Internet Appendix A). The left plot shows the percentage of firms among the 20, 30, and 50 largest financial firms, respectively, that mention economic sanctions-related risk (or cost or uncertainty) in their 10-K reports over the 2015-2023 period. The right plot shows the top eight challenges in sanctions compliance for the 50 largest financial firms.

Similar detection challenges have also been flagged by leading intermediaries in the logistics sector. For example, Daniel Tadros, Chief Operating Officer of the American Club — a leading U.S.-based P&I insurer — was quoted by the New York Times: "It's impossible for us to know on a daily basis exactly what every ship is doing, where it's going, what it's carrying, or who its owners are."

To assess whether information or incentives are the binding constraint on sanctions enforcement, we first assess the credibility of market participants' detection concerns. The ideal design would be a sudden public release of a comprehensive violator list, enabling comparison of market reactions for listed entities against a matched set of compliant tankers. We approximate this setting using a quasi-natural experiment: the unexpected disclosure in August 2023, on the widely used Refinitiv Eikon platform (LSEG, 2023), of a list of suspect tankers compiled by Windward, an AI-based maritime anomaly-detection firm. Following the disclosure, Windward's suspect-tanker module was made publicly available —- free for six weeks and then via a nominal £270/month subscription.

If market participants already had sufficient incentives to avoid violators but were constrained by inadequate information, then by relaxing this constraint the Refinitiv disclosure would have triggered sizable market-driven penalties for sanctions violators. To formally test this hypothesis, we estimate the information elasticity of enforcement: the extent to which improvement in detection accuracy reduces profits for violators.

The first step in estimating this elasticity is to quantify the improvement in detection accuracy following the disclosure. This, however, poses a challenge – the accuracy of interest concerns the probability with which an average sanctions-violating tanker can be identified, whereas we, as econometricians, do not observe the set of true violators. We address this challenge using two complementary strategies. First, we compute the out-of-sample pseudo-$R^2$ with which a third-party could have predicted which tankers would be sanctioned over the subsequent months (until the end of our sample); this enables us, via a structural model, to calibrate the object of interest – the detection accuracy across the broader population of violators. To calculate this out-of-sample pseudo-$R^2$, we train machine learning models using publicly available ship-tracking signals and satellite data, adhering to industry best practices. In the pre-disclosure period, these models achieve a pseudo-$R^2$ of just 4.4%; but after the disclosure, when we include Windward's proprietary list as an additional input, the pseudo-$R^2$ rises to 15.4%. Second, to further validate the informational value of the disclosure, we leverage a unique dataset from an anonymous port agent operating in a Persian Gulf country. This dataset identifies 33 foreign-flagged (non-Iranian) oil tankers that were seen violating Iranian sanctions in January 2021.[2] Of these, 27 were flagged as risky by Windward. By contrast, our machine learning models based solely on public data correctly identified at most 17 of these tankers, while a different but also well-known public list — from the advocacy group United Against Nuclear Iran (UANI) — included only 10.

The second step in estimating the information elasticity of enforcement is to quantify the disclosure's impact on violators' profits. We do so using an event-study framework complemented by structural modeling. For the event study, we implement both a Propensity Score Matched Difference-in-Differences (PSM-DiD) estimator and the semi-parametric Difference-in-Differences approach of Abadie (2005). We compare three key market outcomes for tankers flagged as high-risk by Windward with a matched set of low-risk tankers, before and after the disclosure event. We find that, first, average earnings for high-risk tankers declined by 13% following the disclosure, relative to their counterfactuals. Second, these vessels became 17% less likely to approach sanctioned countries, implying that the perceived risk of detection and enforcement began to outweigh the profit incentives for continued violations. Third, transaction data from Drewry show that the probability of reselling a tanker flagged as risky fell by 37% within a year of the disclosure — likely reflecting restrictions imposed by sanctioning authorities that prohibit transfers of ownership of implicated vessels.

Taken together, these findings point to economically significant losses for suspected sanctions violators following the Refinitiv disclosure. It is important to note, however, that this

[2]While this dataset permits a direct measurement of detection accuracy, we refrain from using it as our main estimate given its limited size and scope.

evidence pertains to *suspected* violators – those flagged on the Refinitiv list, rather than *actual* violators – those targeted by authorities, but whose identities are unobserved. To address this limitation, we develop a dynamic structural model of the oil shipping sector with endogenous compliance decisions. In this framework, key unobservables – such as the accuracy of detecting actual violators, as well as freight rates and profit margins for both sanctions-violating tankers and sanctioned exporters – are determined in equilibrium and can be inferred by calibrating the model to market data and our reduced-form event-study estimates.

Our results indicate that the disclosure increased the probability of correctly identifying actual sanctions-violating tankers from 18% to 42%, triggering a sizable redistribution of rents – about $1.1 billion annually away from violators and $1.9 billion annually toward compliant actors. Sanctioned oil exporters faced a 1.2% increase in freight rates, and their tanker usage declined by 7.7% after the disclosure, while exporters of non-sanctioned oil saw their freight costs drop by 3.5%. Additionally, the continuation value of violating tankers fell by 8.5%, i.e., $3.4 million each for tankers about 15-year-old, similar to those typically used in sanctioned trades. Assuming that roughly 1,500 tankers were engaged in violations but were not formally designated around the time of the Windward event, this translates to a total decline in value exceeding $5.1 billion for these tankers.

In sum, improving detection accuracy imposes substantial losses on sanctions violators, suggesting that better information meaningfully strengthens enforcement. However, our analysis also uncovers a counterintuitive implication: not all compliant actors benefit. Model calibration reveals that even fully compliant tankers – those with no history of transporting sanctioned oil – suffered a decline in earnings following the Refinitiv disclosure, which amounted to an aggregate annual loss of $2.4 billion. This occurs because sanctioned oil exporters, facing heightened detection risk, avoid hiring tankers with prior violations. These tankers shift to the compliant segment of the market, increasing competition and putting downward pressure on freight rates. Part of this downward pressure is offset by sanctioned oil exporters offering a premium to attract compliant tankers, but these added earnings are insufficient to offset the broader decline of freight rates in the much larger compliant market. We test this prediction in Section 3.5 using a shift-share design, and find supportive evidence in the data.

Next we use the model to explore a complementary counterfactual. While the previous analysis assessed improvements in information while holding incentives fixed, now we consider the reverse: strengthened compliance incentives while keeping detection accuracy fixed at pre-disclosure levels. This exercise allows us to test whether limited information was indeed the binding constraint on enforcement. If so, tightening incentives – for example, by imposing

higher fines on exporters of non-sanctioned oil who hire sanctions-violating tankers – should have limited impact in the absence of better detection. The results confirm this hypothesis: with detection remaining as noisy as pre-disclosure, increasing penalties yields little change in enforcement outcomes.

So far, our results show that limited information – rather than weak compliance incentives – was the primary constraint on sanctions enforcement. In order to understand the policy implications of these findings, however, it is important to recognize that this constraint did not stem from a lack of available information: Windward had been selling the same list of sanctions violators prior to its public disclosure, yet most market participants chose not to purchase it. This implies that the market perceived the expected benefit of acquiring the information to be lower than its cost – a fact that has to be accounted for when designing policy.

The simplest explanation for this choice would be that market participants, e.g., exporters of non-sanctioned oil, accurately anticipated the list's value (which we estimate at $4.9 billion for this group) but found its cost to be even higher. However, the data contradict this interpretation. Windward's total stock market valuation – reflecting the value of its sanctions risk list, plus its other assets – was just $42 million prior to the disclosure, suggesting that the cost of acquiring the information was two orders of magnitude smaller than its ex-post benefit. In fact, based on our estimates of total gains and market shares, *each one* of the top 20 non-sanctioned oil exporters gained more than $42 million from the disclosure (Internet Appendix Table A-8), implying that any of them could have acquired Windward outright, released the list publicly, and still profited. It is therefore implausible that our findings reflect a free-rider equilibrium – one where individual exporters lacked sufficient incentive to act and collective action failed due to coordination frictions.

The data, instead, point to a different explanation: exporters of non-sanctioned oil underestimated the accuracy of Windward's list and, as a result, misperceived its value to be low. If the market was stuck in an equilibrium where this misperception persisted—i.e., market forces alone could not correct it—then some form of policy intervention would be warranted.

Such a persistently low valuation for Windward is exactly what we find in the data: Windward's stock price had dropped to half of its IPO value – and stayed at that level or below for over a year – before it decided to disclose its sanctions list on Refinitiv. Our model provides a micro-foundation for why this could happen. At its core is a self-fulfilling feedback loop between beliefs about data accuracy and Windward's market valuation. In this loop, a potential buyer — uncertain about Windward's accuracy, given that it was a young startup in a novel domain (Cohen et al., 2013) — tries to infer what others believe,

using Windward's valuation as a signal (e.g., Sun, 2017; Banerjee, 2011; Bond et al., 2012). A low valuation signals low perceived accuracy, which reduces the buyer's own willingness to pay, thereby reinforcing the low valuation. In such an equilibrium, even a temporary drop in valuation— driven, e.g., by uninformed trading (Edmans et al., 2012)—can lead to persistent undervaluation and overly pessimistic beliefs about accuracy. Windward's decision to release its suspect tanker list at near-zero cost can be interpreted as an attempt to escape this undervaluation trap by publicizing its true accuracy. We present three features of the data consistent with this interpretation in Section 3.2.4.

Even so, one might still ask whether any policy response is necessary—after all, the information constraint was eventually alleviated by the Refinitiv disclosure, without regulatory intervention. However, if that disclosure occurred only as a last resort to escape persistent undervaluation, it raises a deeper concern: would Windward—or any similar firm—have invested in creating such a list if it had anticipated this outcome from the outset? More broadly, the possibility that markets can persistently undervalue even reasonably accurate sanctions-violator lists weakens incentives for private firms to produce them.

Market failure arising from persistent undervaluation of high-quality information implies that authorities need to reconsider information policy as a tool to reduce the profitability of sanctions evasion. Short of publicly funding the production and release of such data, one option is to help private providers credibly signal information quality—for example, through certification mechanisms. Another is to facilitate public access to high-quality data via transfer schemes: for instance, a policymaker could purchase the information from a data vendor at a price that sustains future production—even if that price exceeds what the market would otherwise bear—and then recover the cost by charging a fee or tax to its beneficiaries.

*Related Literature:* Our paper contributes to several strands of literature. First, we relate closely to work on sanctions violation and forensic detection. Prior research has developed creative methods for uncovering violations using indirect signals—such as price gaps in oil markets (Hsieh and Moretti, 2006), stock returns in response to arms embargoes (Dellavigna and La Ferrara, 2010), and rerouted trade during geopolitical blockades (Fisman et al., 2024). Unlike these studies, we examine how markets respond to improved detection, leveraging a quasi-experimental disclosure and a structural model to trace effects on pricing, routing, and asset values. In doing so, we contribute to a broader literature on forensic economics and illegal activity detection (e.g., Dimmock and Gerken, 2012; Fisman and Wei, 2004, 2009; Griffin and Kruger, 2023; Griffin and Maturana, 2016; Griffin and Shams, 2018), while emphasizing information frictions as a structural bottleneck to enforcement.

More broadly, we contribute to the literature on the economics of sanctions and embargoes (e.g., Baldwin, 2020; Cipriani et al., 2023; Crozet et al., 2021; Early and Preble, 2019;

Eaton and Engers, 1992, 1999; Efing et al., 2023; Huynh et al., 2023; Itskhoki and Mukhin, 2023). These papers typically focus on strategic interactions, incentive compatibility, or political economy considerations, and often assume that violations are observable or formally designated. In contrast, we show that improved detection alone can meaningfully enhance enforcement. More specifically, we relate to studies on the effects of recent sanctions on Iran (Dizaji and van Bergeijk, 2013; Haidar, 2017) and Russia (Babina et al., 2023; Huynh et al., 2023; Lastauskas et al., 2023), but differ in focusing on the provision of identifying information about suspected violators, rather than the imposition of sanctions per se. Similarly, while Bai et al. (2025) estimate the global macroeconomic effects of unauthorized shipments sanctioned oil, we highlight the detection limits that allow such shipments to persist despite regulatory efforts.

Finally, we contribute to a broader literature on information disclosure and market discipline (e.g., Goldstein and Yang, 2019; Liberti, Seru and Vig, 2016; Peress, 2014), which examines how public and private signals affect contracting and resource allocation. We show that public classification of suspected violators in enforcement-sensitive markets not only shifts prices but also alters equilibrium dynamics — spillovers that our model formalizes and quantifies.

Taken together, our paper offers the first unified empirical and structural analysis of sanctions enforcement under imperfect detection. We move beyond traditional concerns of incentive compatibility and strategic signaling to show that informational frictions alone can significantly shape enforcement outcomes.

The rest of the paper is organized as follows: Section 1 provides a brief background, Section 2 examines the Refinitiv disclosure, Section 3 presents a dynamic structural model and estimates key quantities of interest within its scope, and Section 4 concludes.

# 1 Background

In this section, we first explain why it may be difficult to detect violating tankers – in spite of their large, slow-moving nature, and in spite of modern advances in data and technology to monitor vessels at sea. Then we describe the circumstances surrounding the key event we study – the Refinitiv disclosure in August 2023 – that allows us to evaluate whether a lack of information or incentives is acting as a binding constraint on improved sanctions enforcement.

## 1.1 Obfuscating Violations: The Case of Oil Transport

Before we focus on detecting sanctions-violating tankers, here we explain what makes detection hard – i.e., how do violating tankers obfuscate their activities. Here we illustrate typical strategies violators employ to evade detection through ship-tracking data from the Automatic Identification System (AIS).[3]

Authorities enforcing sanctions advise the shipping industry (OFAC, 2023) to monitor suspicious AIS transmission gaps ("dark activity"), and geolocation "spoofing", where AIS data is falsified to display a ship in a false location, akin to using a VPN (Windward, 2020). Such spoofing, reported by the UN in 2019 (U.N. Doc. S/2019/171), has rapidly proliferated since 2021 (New York Times, Sep. 2022, Economist, Apr. 2022).

Figure 2 illustrates spoofing and dark activity for two tankers in our Iranian sanctions violators dataset. In the top panel, the tanker's AIS signals falsely indicated its presence in the northern Persian Gulf while it was actually in Iran according to our ground truth data. The tanker in the bottom panel ceased AIS transmissions for four days during which it was also observed in Iran.

Regulators (OFAC, 2020), and industry experts (e.g., Wolsing et al. (2022)) suggest multiple categories of sanctions violation predictors to detect such activity: (i) tanker identity change, (ii) using risky country flags, (iii) ship-to-ship transfers, (iv) irregular trajectories (that are thought to indicate falsified routes), (v) anomalous locations (where tankers typically do not tread), and (vi) dark activity. In addition, more sophisticated players are also advised to use satellite data for detection (OFAC, 2023, Economist, Apr. 2022), e.g., from the Sentinel-1 mission. Figure A-1 in the Internet Appendix displays three of these images as an example, with further accompanying details.

We show in the next Section that detection remains highly challenging, even if this advice is followed diligently by compliant parties.

## 1.2 The Refinitiv disclosure

On 16 August 2023, the maritime AI firm Windward made its sanctions risk classifications for oil tankers publicly accessible via the London Stock Exchange Group's (LSEG) Refinitiv Eikon. This data platform, originally started by Reuters, and Bloomberg are the two most widely used business information platforms in the world – Bloomberg has a 33.4% market share, while Refinitiv Eikon has 19.6% across 190 countries worldwide (Investopedia).

[3] AIS was originally designed for collision avoidance. Vessels over 300 gross tonnage must carry AIS equipment (IMO, 2000). This data has been utilized in academic research before (e.g., Brancaccio et al. (2020)), is also used for sanctions compliance, (U.N. Doc. S/2019/171, Kilpatrick (2022)), and even flag state registries have used it to de-flag ships (Lloyd's List, Feb. 2020, Oct. 2020).

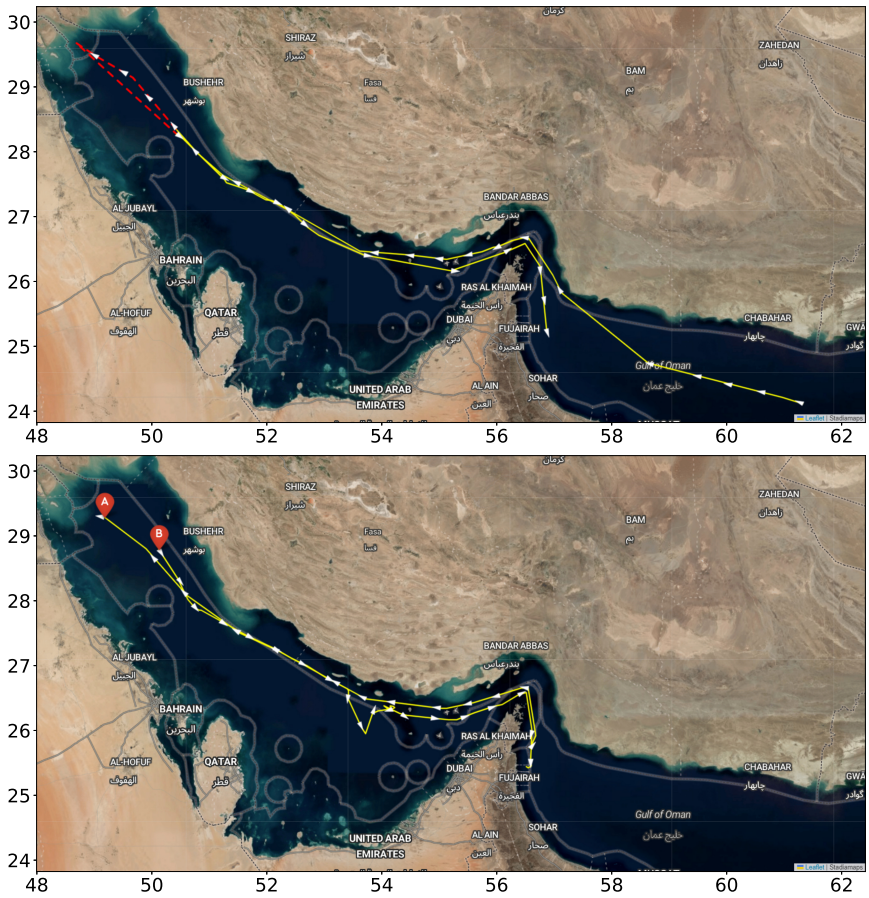



Figure 2: In this figure, solid and dashed lines show the trajectories of two tankers in the Persian Gulf, as given by their AIS signals. White arrows indicate the direction of movement. The top plot shows a case of spoofing, where a tanker, which is placed at an Iranian port by our dataset of sanctions-violating tankers, at the same time emits falsified AIS signals (red dashed lines) showing it traveling in the northern part of the Gulf near Iraq. The yellow solid lines show this tanker's path before and after its Iranian port visit. The bottom plot shows a case of dark activity, i.e., the tanker stopped emitting signals while observed at an Iranian port (per our dataset). Marker "A" indicates the last signal before it went dark, emitted at 02:16:12 on 2021-01-23. Marker "B" shows the first signal after its dark period, emitted at 04:54:45 on 2021-01-27 not too far from A and near the Iranian coast, concluding a four-day dark period.

The release therefore marked a notable shift in the accessibility of sanctions risk information. The list itself categorizes over 400,000 vessels into Low, Moderate, High, and Sanctioned risk tiers, leveraging over 100 million daily data points, integrating proprietary satellite imagery and weather data with information from its clients, which include the United Nations, the European Border and Coast Guard Agency (Frontex), and U.S. agencies such as the Drug Enforcement Administration and the Office of Naval Intelligence (Reuters, Mar. 2016, Wired, Mar. 2020). Observers have compared Windward's technology to military-grade signals intelligence adapted for commercial use (RAND, 2017). The Windward list had previously only been bought by a limited set of clients. but by publishing the data on

Refinitiv Eikon — where it was available for free for six weeks, followed by a subscription fee of approximately £270/month — this disclosure effectively democratized access to advanced maritime risk intelligence.

To understand what might have motivated Windward to disclose this data for free – an issue which we study in detail later – we describe here what was happening with the company around that time. Windward was the first pure-play maritime data company to publicly list on an exchange, with a market capitalization of approximately £126.5 mln in December 2021. But by the second week of August 2023, right before it disclosed the list on Refinitiv, this value had steadily declined to only £35 mln (about $42 mln; see, e.g., Yahoo Finance, 2023).

However, in spite of publicly releasing the sanctions suspect list initially for free, the disclosure was followed by a 10% increase in Windward's stock price over the next 3 trading days. This is likely because the disclosure drew attention to, and boosted confidence in Windward's capabilities in sanctions compliance AI and, broadly, in maritime AI. This increased overall sales for Windward by demonstrating the company's capabilities and helping secure new contracts – following the announcement, Windward shifted its focus toward offering tailored solutions for both commercial and government clients (Proactive, 2023). The stock gained further momentum in the months following; by March 2024, Windward's market cap had increased to around £179 mln. Finally, in December 2024, Windward announced it would go private after agreeing to a takeover bid by U.S.-based growth equity firm FTV Capital. The acquisition valued Windward at £216 mln.

# 2 Information shocks and market-based enforcement

## 2.1 The disclosure's effect on the difficulty of detecting violators

We first assess whether Windward's list helps third parties predict which tankers are about to be sanctioned in the near future. Such predictive ability is crucial for third parties' risk management; e.g., if a bank ended up unknowingly lending to such a tanker, the loan amount might be at risk if the tanker got sanctioned.

To do so, we first take the view of a third-party that uses AIS-signal-based ship-tracking data, as well as satellite imagery, as recommended by authorities in Section 1. We assume that these third parties use these data to estimate Machine Learning (ML) models, which allow for the most flexibility (e.g., various non-linear combinations of the above predictors), and are designed for such prediction exercises. Specifically, we employ decision trees and neural networks. We use 10-fold cross-validation to fine-tune model parameters; the sample

is randomly divided into 10 sub-samples, models are trained with nine sub-samples, leaving one for validation. This procedure is repeated 10 times to select the best hyper-parameter combination based on McFadden's pseudo-$R^2$. Detection performance is evaluated from the best cross-validated model using its pseudo-$R^2$.

Our goal here is to predict which tankers will be sanctioned between August 2023 and May 2024 (the end of our sample), using data up to July 2023. First, we train ML models based on the sample from August 2022 to July 2023 using tankers sanctioned during the training period as our targets (i.e., we exclude tankers sanctioned before August 2022). Then, for out-of-sample prediction, we focus on the tankers that have not yet been sanctioned until July 2023, and average model predictions from a decision tree and a neural network. The predictors, including both AIS-based and satellite-based indicators, are described in Internet Appendix B.

To compare Windward's predictive ability, a simple metric is to rank tankers by our model-predicted probabilities and select the top 25% – the same proportion as Windward, and then compare the number of tankers correctly predicted to be subsequently sanctioned. Among 69 tankers that are sanctioned post-July 2023, our list of the top-25% only contains 27, while Windward's list contains 65, demonstrating its substantially superior accuracy. More formally, as we show in Panel A of Table 1, the pseudo-$R^2$ with which to-be-sanctioned tankers can be predicted by ML models using public data alone (i.e., AIS+satellite) is 4.4%. In contrast, optimally combining Windward with the public data (0.55 × public data + 0.45 × zero-one dummy of Windward's risk label) raises pseudo-$R^2$ to 15.4%.[4] Panel B of the Table shows the importance of individual predictor variables in terms of the drop in pseudo-$R^2$ when each is left out – and again reveals that Windward's list is by far the most important of these.

Next, we benchmark Windward's relative accuracy against AIS-based alternatives using a "ground truth" dataset on sanctions-violating oil tankers in Iran. This unique dataset lists all non-Iranian-flagged tankers seen in Iran's Persian Gulf waters in January 2021 (total of 33 tankers), and is collected by an anonymous source in the Middle-Eastern shipping industry via text messages from port agents, brokers, and charterers. While the provenance and scale of the dataset limit its usefulness, it offers a rare opportunity to test detection models against a set of *confirmed* breaches – something missing in the literature (Wolsing et al. (2022)).[5]

Internet Appendix Table A-2 Panel C shows that Windward classified 27 out of these 33 listed violators as high/moderate risk, whereas our models detected only 17 at best, even at

[4] Figure A-6 in the Internet Appendix shows how different weights on these two approaches affect overall pseudo-$R^2$s, in particular that public data-based ML still has value even after the Refinitiv disclosure.

[5] We intend to make this dataset publicly available to facilitate replication and further studies.

Table 1: **Out-of-sample predictability of future sanctioned tankers**

This table reports the out-of-sample predictability of future sanctioned tankers based on public data (AIS data and satellite images) or public data combined with Windward risk labels. The training sample spans from August 2022 to July 2023, and the testing sample spans from August 2023 to May 2024 (end of our sample). In panel A, we report the model performance measured by the out-of-sample pseudo-$R^2$. For the column "AIS + Satellite", we average model predictions from a decision tree and a neural network. For the column "AIS + Satellite + Windward", we use the optimal linear combination between public signal and Windward risk labels (a zero-one dummy). In panel B, we report the predictor importance measured by the reduction of the out-of-sample pseudo-$R^2$ when setting the corresponding predictor(s) to zero.

| Panel A: Model Performance | | |
|---|---|---|
| | AIS + Satellite | AIS + Satellite + Windward |
| pseudo-$R^2$ | 4.35% | 15.38% |
| **Panel B: Predictor Importance** | | |
| Windward | | 11.03% |
| Satellite Detection | 2.69% | 1.05% |
| Identity Change | 0.03% | 0.00% |
| Risky Flag | 0.57% | 0.18% |
| Irregular Trajectory | 0.66% | 0.33% |
| Ship-to-ship Transfer | 2.67% | 1.06% |
| DBSCAN Outlier | 1.20% | 1.03% |
| Dark Activity | 0.00% | 0.87% |

a 90% confidence level. While these results are not directly comparable – since Windward may have flagged tankers based on activity beyond January 2021 – its significantly higher detection rate is notable. Moreover, even UANI's "The Ghost Armada" list – also based on activity not restricted to January 2021 – recognized just 10 of the 33 violators at the time of the Refinitiv disclosure, further underscoring Windward's superior accuracy in detecting sanctions violators.

Note that Bai et al. (2025) also aim to detect sanctions-violating oil tankers and report higher predictive accuracy – an 83.3% overlap between their model's predictions and a set of target vessels. However, their targets are not confirmed violators, but a small subset of 25 suspected violators publicly named by commercial sources (e.g., Lloyd's List, S&P Global Market Intelligence).[6] Their reported accuracy could be higher than ours because (i) the tankers they target – those already detected by Lloyd's List and others as suspects, and in particular those publicly flagged – are likely to be more salient and easily detectable, overstating accuracy and (ii) their accuracy metric ignores false positives, assessing only the proportion of their chosen subset that their detection algorithm successfully predicts. In contrast, we (i) aim to detect *actual* violators – either tankers later officially

[6]These 25 tankers (with 42 tanker-year observations) represent less than 10% of the full set of suspects listed by these sources (e.g., one of the Lloyd's list reports that they cite mentions a total of 450 suspects).

sanctioned or observed violating Iranian sanctions in our ground truth data; and (ii) use McFadden's pseudo-$R^2$, a standard metric that penalizes false positives. When we use their list of suspected violators as the target and their definition of accuracy, our models achieve performance comparable to Bai et al. (2025) (see Internet Appendix D.4), indicating that differences in reported accuracy mainly reflect differences in targets and metrics rather than modeling strength.

Finally, we find that market participants were aware of, and attentive to, the disclosure: Figure A-5 in the Internet Appendix shows how Windward's homepage views saw a sharp increase following the list's release.

## 2.2 Real effects of the disclosure on suspect tankers

Here we examine the real effects of the Refinitiv disclosure on suspect ships, focusing on (i) tanker fixtures (i.e., freight rates or rental rates negotiated between a shipowner and a charterer), (ii) tanker route changes, and (iii) ownership transitions.

We start by estimating the effect of Windward's high/moderate risk classification on affected tankers' fixtures in a matched-sample Differences-in-differences (DiD) framework. Fixtures are measured in standardized Worldscale units, allowing for easy comparison across contracts (see Internet Appendix E.3 for details). We examine the change in fixtures around the disclosure for tankers labeled as high/moderate risk, and compare them to counterfactual tankers – otherwise identical ones, but labeled by Windward as low risk. The challenge lies in constructing such counterfactuals, and we use two methods to do so.

First, we use propensity score matching (PSM), introduced by Rosenbaum and Rubin (1983); this method looks to match a treated, i.e., a high/moderate risk tanker, to a low risk one which had the same ex-ante propensity of being classified as high/moderate risk based on public data. We present summary statistics and balance tests – which show that the matched tankers are ex-ante very similar to the treated ones – in Internet Appendix Table A-4. This method is intuitive, and allows for standard DiD plot visualization.

In this framework, treatment effects for high-risk tankers are estimated as follows:

$$fixture_{c,i,k,t} = \left[ \sum_{l=-6}^{7} \beta_l^H \times high_risk_i \times \mathbb{I}_{\{t=l\}} \right] + \alpha_i + \gamma_{k,t} + \epsilon_{c,i,k,t}, \tag{1}$$

where $fixture_{c,i,k,t}$ is the WS rate for contract $c$, tanker $i$ (with tanker type $k$) in month $t$; $high_risk_i$ equals one for tankers labeled as high-risk in Windward's list, and zero otherwise; $\alpha_i$ is tanker fixed effects, and $\gamma_{k,t}$ is time×tanker-type fixed effects. We are interested in the average treatment effect on the treated (ATT) for high-risk tankers, given by the coefficient

series $\{\beta_l^H\}_{l=-6}^{7}$. To keep interpretation simple, we first exclude tankers classified as moderate risk from the sample.

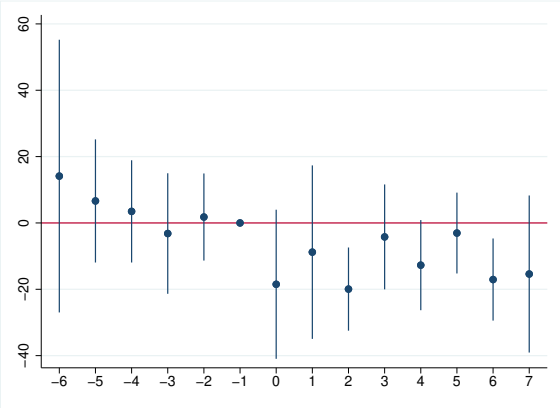
Plot A in Figure 3 shows that the PSM-DiD estimated coefficients in the pre-period $\{\beta_l^H\}_{l=-6}^{-1}$ are insignificant, indicating a lack of pre-trends. Post-disclosure, fixture rates for high-risk tankers drop immediately in August by about 20 WS units compared to the matched control group, a change that persists until the end of our sample in March 2024. To assess the impact on tanker earnings, consider that the average fixture for high-risk tankers in July 2023 was 126.8 WS units (Internet Appendix Table A-4). A 20-unit drop in August implies a 15.8% revenue decrease (20/126.8). Note that this decrease reflects the market reaction to the Windward-induced change in the risk of tankers violating sanctions, not actual violations.

Next, in Plot B of Figure 3 we examine the impact of the information shock on tanker routes, i.e., the propensity of newly labeled high/moderate risk tankers to show up in Iranian, Russian, or Venezuelan territorial seas (defined using data from the Flanders Marine Institute). The dependent variable is a zero-one indicator of AIS signals emitted within the territorial seas of these countries each month. Our test design mirrors the fixtures analysis from the previous section. This plot shows estimated coefficients $\{\beta_l^H\}_{l=-6}^{-1}$, which show probabilities of passing near sanctioned countries. Pre-period coefficients show a lack of pre-trends. Post-period, there is a gradual drop reaching about 10 percentage points at the end of our sample.

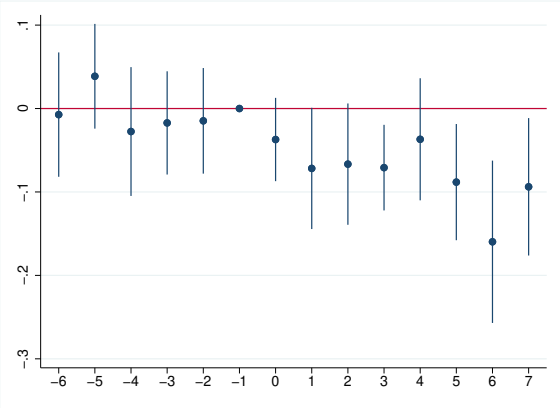
Finally, in Plot C of Figure 3 we examine whether tankers newly classified as high/moderate-risk become harder to sell after the disclosure. This hypothesis is based on the fact that many sanctions regimes explicitly prohibit transactions involving violating entities (e.g., see Council Regulation (EU) 833/2014 on such restrictions, Skadden, 2024). Here, we use proprietary tanker ownership data from Drewry, showing owner names and countries in six snapshots between December 2022 and March 2024. The dependent variable equals one if owner names differ across two snapshots and zero otherwise. To account for different time spans between snapshots, we annualize the ownership change variable. Plot C shows that the disclosure has indeed reduced the turnover of high/moderate-risk tankers (which we merge together in this panel due to small sample sizes here), making them less desirable.

While this visual evidence is suggestive, and post-period coefficients in Figure 3 are uniformly lower than the pre-period ones, month-by-month estimation is noisy. To mitigate noise, we aggregate post-period indicators into a single dummy $\mathbb{I}_{\{t\geq 0\}}$, and estimate ATT separately for high-risk and moderate-risk tankers. In a matched sample with high- and

A. Change in fixtures




B. Change in routes




C. Change in ownership

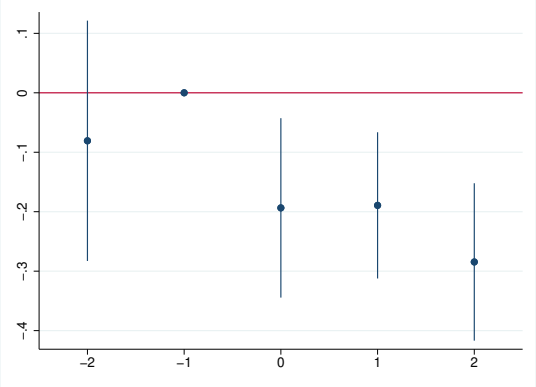



Figure 3: Plot A. shows the time series of coefficients (i.e., average treatment effects on the treated, ATT) estimated with PSM-DiD to examine the change in fixtures for high-risk tankers from Feb 2023 to Mar 2024 (end of our fixture data). 0 indicates the Refinitiv disclosure in Aug 2023. The propensity score is calculated by regressing the high-risk tanker indicator on pre-period tanker characteristics using ML methods: average outputs from a decision tree and neural networks, with variables described in the Internet Appendix B. We match tankers within each tanker type and calculate weights based on propensity scores and a Gaussian kernel with a bandwidth of 0.01. Then we regress fixtures (trimmed at the 1st and 99th percentiles) on indicators of high-risk tankers interacted with period dummies, controlling for tanker fixed effects and time×tanker-type fixed effects. 95% confidence intervals using standard errors double clustered by tanker and time×tanker-type levels are shown. Plot B. shows analogous coefficients to examine the routes taken by high-risk tankers. The dependent variable equals one in a month when a tanker passes within 12 nautical miles of Iran, Russia, or Venezuela (only before 2024) from Feb 2023 to Mar 2024. Plot C. shows analogous coefficients to examine the effect of the disclosure on the probability of tanker owner changes for high-risk tankers. Our proprietary tanker ownership data contains six snapshots: Dec 2022, Mar, Jul, Sep, Dec 2023, and Mar 2024. The dependent variable is a zero-one indicator of owner changes by comparing the owner names across two snapshots. Since the time spans between consecutive snapshots are different, we make them comparable by annualizing the owner change variable. The pre-period includes Dec 2022 - Mar 2023 and Mar - Jul 2023. The post-period includes Jul - Sep 2023, Sep - Dec 2023, and Dec 2023 - Mar 2024.

low-risk tankers, we run the following regression:

$$fixture_{c,i,k,t} = \beta^H \times high_risk_i \times \mathbb{I}_{\{t \geq 0\}} + \alpha_i + \gamma_{k,t} + \epsilon_{c,i,k,t}, \tag{2}$$

where $\beta^H$ is the ATT for high-risk tankers. We perform a similar estimation for moderate-risk tankers by replacing $high_risk_i$ with $moderate_risk_i$. The first three columns in Panel A of Table 2 presents the PSM-DiD regression results for fixtures.

Next, we check for the robustness of our findings to using a different DiD estimator, as in Abadie (2005). Compared with ad hoc PSM-then-DiD implementations, Abadie (2005) provides a semiparametric matching DiD estimator that identifies the ATT under parallel trends in untreated potential outcomes and allows for time-invariant unobserved heterogeneity. The estimator reduces reliance on parametric outcome specifications and offers a more standardized procedure with established asymptotics. Further details are in Internet Appendix E.

Since Abadie (2005) is a matching-based DiD estimator and does not include regression-style fixed effects, we pre-process outcomes to net out additive macro trends. Specifically, for each month t we subtract the cross-sectional mean outcome among untreated fixtures, and in an alternative specification we subtract the month-by–tanker-type mean (and thus adjust for tanker-type-level monthly changes in freight rates) among untreated fixtures. We then implement the Abadie (2005) matching estimator on first-differenced outcomes. Results are robust to omitting this pre-processing, and we report all three versions of the estimated ATT, i.e., using raw fixtures, fixtures demeaned by time, and fixtures demeaned by time and tanker type, respectively.

The first three columns in Panel B of Table 2 shows results from this method for fixtures. The sample size in panel B is smaller than in panel A, because Abadie (2005)'s estimation uses a tanker-level sample (not a tanker-month level one), and also excludes tankers without both pre- and post-period fixtures (because the dependent variable is post-period average fixture minus pre-period average fixture).

The results show that, after the information shock, the fixtures of high-risk tankers decreased by an average of 16.45 WS units, ranging from 13.75 to 20.71 units. This translates to a 13% earnings drop (16.45/126.8), relative to the counterfactual. No significant change is observed for moderate-risk tankers.

Of note here is that while matching on observables reduces imbalance, we recognize that unobserved differences between risky and low-risk tankers may still bias our estimates. To mitigate such concerns, we assess the robustness of our results. Our evidence suggests that they remain similar, not only across multiple methods and bandwidth choices, but also when

we use logistic regression for propensity scores, drop August 2023 (to allow about two weeks for the disclosure effect to come into force), omit Russia, and use bootstrap standard errors. We present these robustness tests in Table A-5 in the Internet Appendix.

The middle three columns in Table 2 present the estimation of ATT on route changes, combining post-period months into a single dummy, as we did for fixtures. As AIS signals are observed much more frequently than fixtures, this route test uses more observations.

The PSM-DiD results in Panel A show that high-risk tankers are about seven percentage points less likely to pass close to sanctioned countries, with no significant effect on moderate-risk tankers. The average probability of signals in these areas is 40 percentage points for high-risk tankers in July, it drops by 17% (6.8/40) in the post-period. Panel B using Abadie (2005)'s semiparametric DiD yields similar results. Overall, our evidence suggests that high-risk tankers avoid approaching sanctioned countries after the disclosure.

The last three columns in Table 2 present results on tanker ownership changes in a regression setting similar to the other columns. Our evidence shows that that after the information shock, high-risk (moderate-risk) tankers are 16.4-18.5 (14.6-16.8) percentage points less likely to change owners. This represents a 37% drop in liquidity for high-risk tankers, relative to a pre-disclosure annualized turnover of 45.4 percentage points (17/45.4).

In one further exercise (reported in Internet Appendix Table A-6) we examine whether U.S. and U.S.-allied charterers start avoiding tankers classified as high or moderate risk by Windward. Our evidence shows a significant drop in high-risk tankers usage by U.S. charterers, while effects for U.S.-allied charterers are not statistically significant.

Finally, we examine whether selection issues affect our data, specifically if the information shock altered high/moderate-risk tankers' reporting of their fixture contracts to our data vendor. Results in Internet Appendix Table A-7 show no evidence of changes in reporting by risky tankers. Note that our ownership change data does not come from reports to Refinitiv, and is not affected by such potential selection issues.

# 3 Sanctions and Disclosure through the Lens of a Model

In this section, we develop a dynamic structural model of intermediaries and sanctions, and calibrate it using data from the oil shipping market and our Windward event-study results from the preceding section. The model (i) enables us to quantify the impact of the Refinitiv disclosure on two key actors for whom data is unavailable – actual sanctions-violating tankers (as opposed to suspected violators detected by Windward) and exporters of sanctioned oil. Additionally, (ii) the model suggests one way to reconcile Windward's ex-ante low valuation with the substantial changes its disclosure produces. This particular

Table 2: **Difference-in-differences analysis**

This table reports the coefficients (i.e., average treatment effects on the treated, ATT) estimated with PSM-DiD and Abadie (2005)'s semiparametric DiD to examine the change in (i) fixtures, (ii) routes, and (iii) ownership for high/moderate-risk tankers after the Refinitiv disclosure. The sample period is from Feb 2023 to Mar 2024, and the post-period starts from Aug 2023. In panel A, we match tankers within each type and calculate weights based on propensity scores and a Gaussian kernel with a bandwidth of 0.01, 0.03, or 0.05. Then, we do DiD estimation in the matched sample, controlling for tanker fixed effects and time×tanker-type fixed effects, with standard errors double clustered at the tanker and time×tanker-type levels. In panel B, we implement Abadie's method by collapsing our sample into two cross-sections by calculating each tanker's pre- and post-period averages of the dependent variables. Since Abadie's estimator is derived without directly accounting for macro trends, we manually subtract the cross-sectional mean by tanker type to account for time×tanker-type fixed effects. We trim the fixtures at the 1st and 99th percentiles each month to avoid the influence of outliers, and derive standard errors as in Abadie (2005). *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

| Panel A: PSM-DiD | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dependent variable | I. Fixtures | | | II. Showing AIS signals near sanctioned countries | | | III. Owner change | | |
| Bandwidth | 0.01 | 0.03 | 0.05 | 0.01 | 0.03 | 0.05 | 0.01 | 0.03 | 0.05 |
| ATT for High Risk | -16.06*** | -15.15** | -14.61** | -0.073*** | -0.072*** | -0.071*** | -0.183*** | -0.169*** | -0.164*** |
| | [-2.69] | [-2.39] | [-2.31] | [-3.69] | [-3.76] | [-3.79] | [-3.25] | [-2.91] | [-2.87] |
| Obs. (tanker-months) | 5,378 | 5,574 | 5,574 | 39,246 | 39,246 | 39,246 | 11,402 | 11,402 | 11,402 |
| ATT for Moder. Risk | -0.68 | -1.20 | -1.49 | -0.016 | -0.015 | -0.015 | -0.168*** | -0.159** | -0.152** |
| | [-0.17] | [-0.32] | [-0.41] | [-1.01] | [-1.02] | [-1.05] | [-3.30] | [-2.79] | [-2.50] |
| Obs. (tanker-months) | 6,081 | 6,083 | 6,083 | 43,820 | 43,820 | 43,820 | 11,776 | 11,776 | 11,776 |
| Time × Type FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Tanker FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Panel B: Abadie (2005) Semiparametric DiD | | | | | | | | | |
| | raw | by time | by time × type | raw | by time | by time × type | raw | by time | by time × type |
| ATT for High Risk | -18.42*** | -20.71*** | -13.75*** | -0.067*** | -0.067*** | -0.060*** | -0.185*** | -0.184*** | -0.170*** |
| | [-3.56] | [-4.47] | [-3.16] | [-4.01] | [-4.03] | [-3.60] | [-3.30] | [-3.29] | [-3.04] |
| Obs. (tankers) | 1,036 | 1,036 | 1,036 | 2,543 | 2,543 | 2,543 | 2,346 | 2,346 | 2,346 |
| ATT for Moder. Risk | -3.74 | -5.09 | -0.51 | -0.012 | -0.012 | -0.008 | -0.168*** | -0.167*** | -0.146*** |
| | [-0.88] | [-1.40] | [-0.15] | [-0.97] | [-0.96] | [-0.62] | [-3.20] | [-3.18] | [-2.77] |
| Obs. (tankers) | 1,190 | 1,190 | 1,190 | 3,139 | 3,139 | 3,139 | 2,421 | 2,421 | 2,421 |

modeling choice also provides one explanation for why Windward might have been stuck in a low-valuation trap, which might, in turn, have led them to disclose the list at low cost to publicize their accuracy. Finally, (iii) we use the model to generate a counterfactual where sanction penalties are increased further.

## 3.1 Theoretical Framework

### 3.1.1 Environment

We model the oil-shipping sector. Our economy consists of oil exporters and tankers.

- 'Rogue' (R) exporters deal in sanctioned products (e.g., Iranian oil companies), while 'Clean' (C) exporters deal in non-sanctioned, legally-traded products (e.g., U.S. oil companies).

- Tankers are either 'Good' (G), i.e., those that have never carried sanctioned oil, or 'Bad' (B), i.e., those that have done so at least once in the past. A Bad tanker cannot revert back to being Good, reflecting the fact that tankers can be sanctioned for past violations. Once a Bad tanker is sanctioned and put on a designated list it cannot be hired by any exporter. However, some Bad tankers are not sanctioned yet, and these are the main source of risk for an exporter in our model, who faces penalties and disruptions if such tanker is sanctioned while in their employ.

- Third parties like banks, insurers, tanker charterers, etc. who provide intermediary services face similar risks if engaging with Bad tankers. Therefore, we do not explicitly include them in the model for simplicity; one can think of them as all being folded into the Clean exporter category.

- Rogue exporters know which tankers are Bad because they have employed them previously. Clean exporters only know that a proportion $\lambda$ of all tankers are Bad ($\lambda$ is common knowledge). They use a noisy detection technology that classifies a portion $\lambda$ of tankers as High-risk (H), and the rest as Low-risk (L). We assume a stationary environment where the proportion of Bad tankers, $\lambda$, is constant.

- Tankers know if they are Bad, and can also infer their risk label (H or L) from the fixture quoted to them by a randomly matched Clean exporter (we assume, for simplicity, that search frictions prevent tankers from shopping for rates from different Clean exporters). Rogue exporters do not know the fixtures quoted by Clean exporters (who compete with them for tankers), and hence cannot infer risk labels.

### 3.1.2 Timeline

We assume the following sequence of events each period:

1. Clean exporters optimally choose detection technology and assign risk labels to tankers.
2. An aggregate i.i.d. mean-zero shock $\tilde{\epsilon}$ to the oil trading revenue of all exporters realizes, reflecting unpredictable oil price changes that exporters cannot hedge ex-ante.
3. All exporters choose simultaneously freight rates to offer to tankers: Clean exporters choose $p_L$ and $p_H$ for L and H tankers, respectively, and Rogue exporters choose $p_G^R$ and $p_B^R$ for Good and Bad tankers.
4. Each tanker receives two freight rates: one from a randomly matched Clean exporter and another from Rogue exporters, and decides whom to engage with.
5. The sanctioning authority sanctions (a subset of) violators.
6. Each player receives a payoff and proceeds to the next period (tankers continue to the next period only if not sanctioned in this period).

### 3.1.3 Clean exporters' detection technology

**Signals**

There are $N$ Clean exporters, indexed by $i$. To detect whether a given tanker $j$ is Bad, Clean exporters rely on three potential sources of information (we omit here the tanker index $j$ for brevity).

(i) A common signal $s^{C,com}$, based on public information, available at zero cost and with fixed precision $\sigma_\zeta^{-2}$:

$$s^{C,com} = \mathbb{I}(\text{Bad Tanker}) + \zeta, \quad \zeta \sim \mathbb{N}(0, \sigma_\zeta^2), \quad \zeta \perp\!\!\!\perp u. \tag{3}$$

(ii) A private signal $s_i^{C,pri}$, with precision $\sigma_{\xi,i}^{-2}$:

$$s_i^{C,pri} = \mathbb{I}(\text{Bad Tanker}) + \xi_i, \quad \xi_i \sim \mathbb{N}(0, \sigma_{\xi,i}^2). \quad \xi_i \perp\!\!\!\perp u, \zeta, \tag{4}$$

Clean exporter $i$ can reduce $\sigma_{\xi,i}$ by paying a (continuous) cost:

$$\Omega(\sigma_{\xi,i}) = \kappa \sigma_{\xi,i}^{-2}, \tag{5}$$

where $\kappa$ is a constant calibrated by matching moments.[7]

[7] This functional form satisfies the no-arbitrage condition in information acquisition: the cost of independently acquiring two signals with (im)precision $\sigma_A$ and $\sigma_B$ and then optimally combining them, $\Omega(\sigma_A) + \Omega(\sigma_B)$, is the same as the cost of directly acquiring a signal with the combined (im)precision, $\Omega(1/\sqrt{\sigma_A^{-2} + \sigma_B^{-2}})$.

(iii) A signal $s^W$ that can be bought for a fixed price $P^W$ from a specialized data provider like Windward. In our model, such a profit-motivated data intermediary plays a key role, as its data release represents the information shock that we examine.

As Windward integrates information from multiple sources, that include the United Nations, European Border and Coast Guard Agency (Frontex), and various U.S. agencies, among others, we model Windward's signal as a noisy version of a signal $s^A$, which is only available to the sanctioning authority:

$$s^W = s^A + v, \quad v \sim \mathbb{N}(0, \sigma_v^2), \quad v \perp\!\!\!\perp u, \zeta, \xi. \tag{6}$$

$s^A$ determines the probability of sanctioning the given tanker, as specified below.

The common signal in Eq. (3) reflects due diligence conducted using public data. The private signal in Eq. (4) captures the requirement imposed by the sanctioning authority that market participants do their own due diligence, which is motivated by its belief that these participants have private information on suspicious sanctions-violating behavior. Although we lack details on the exact private information Clean exporters have, industry reports reveal the equilibrium cost they pay for due diligence, enabling us to calibrate relevant parameters.

Windward's signal $s^W$ in Eq. (5) – unlike the common signal $s^{C,com}$ – contains information about the authority's signal which determines sanctions. However, Clean exporters do not know the precision of this signal ex-ante, i.e., they do not know $\sigma_v^2$ before the disclosure, but form beliefs about it. In Section 3.2.3, we specify how their beliefs can be calibrated from Windward's observed equity market value. For now, let $\tilde{\sigma}_{v,i}^{-2}$ denote Clean exporter $i$'s posterior belief on Windward's precision.

**From signals to risk labels:**

Clean exporters optimally choose whether to buy Windward's signal, to what extent they want to improve the precision of their private signal by paying the associated cost, and how to combine the signals. For the latter, we assume that they weigh each signal by the inverse of its noise variance.

Let $\chi = 1$ ($\chi = 0$) indicate a Clean exporter's decision to buy (not buy) Windward's information (we study a symmetric equilibrium in which all Clean exporters choose the same action). Before the disclosure, the expected precision of the combined signal is $\sigma_\zeta^{-2} + \sigma_\xi^{-2} + \chi \cdot (\sigma_u^2 + \tilde{\sigma}_v^2)^{-1}$, and the cost is $\Omega(\sigma_\xi) + \chi \cdot P^W$. After the disclosure, clean exporters have free access to $s^W$ and can observe its precision. Accordingly, the combined signal's precision becomes $\sigma_\zeta^{-2} + \sigma_\xi^{-2} + \chi \cdot (\sigma_u^2 + \sigma_v^2)^{-1}$, and the cost reduces to $\Omega(\sigma_\xi)$.

Let $s^C$ denote the optimally combined signal. Clean exporters assign risk labels to tankers by applying a threshold $K$ to $s^C$: a tanker is classified as High-risk (H) if $s^C$ exceeds $K$, and

as Low-risk (L) otherwise. The threshold $K$ is chosen to satisfy

$$\mathbb{E}(s^C \geq K) = \lambda. \tag{7}$$

### 3.1.4 Sanctions

The sanctioning authority allocates to each tanker $j$ monitoring resources $a_j$, which depend on a noisy signal $s_j^A$ that it receives about the tanker's type:

$$a_j = f(s_j^A), \quad \text{and} \quad f'(\cdot) > 0. \tag{8}$$

where

$$s_j^A = \mathbb{I}_j(\text{Bad Tanker}) + u_j, \quad u_j \sim \mathbb{N}(0, \sigma_u^2). \tag{9}$$

To simplify analysis, we specify

$$f(s_j^A) = \frac{\exp(s_j^A)}{\mathbb{E}[\exp(s_j^A)]}, \tag{10}$$

so that $a_j$ is always positive and $\mathbb{E}[a_j] = 1$.

The probability that a given tanker $j$ is sanctioned is:

$$\mathbb{P}_j(\text{sanctioned}) = a_j \cdot [\pi_1 \cdot \mathbb{I}_j(\text{Bad tanker}) + \pi_2 \cdot \mathbb{I}_j(\text{deal with R})], \tag{11}$$

where $\mathbb{I}_j(\text{Bad tanker})$ equals one if this tanker is Bad and zero otherwise, $\mathbb{I}_j(\text{deal with R})$ equals one if the tanker is currently transporting sanctioned oil and zero otherwise, $\pi_1$ is the probability per unit of resource that the tanker is sanctioned for past violations (i.e., for being Bad), and $\pi_2$ is this probability of a current violation (i.e., transporting sanctioned oil from Rogue exporters in the current period). The parameters, $\pi_1$ and $\pi_2$, reflect legal and other technical constraints that the authority faces in sanctioning tankers.

### 3.1.5 Beliefs about exposure to sanctions-related penalties

**Tankers**: Depending on its type and risk label, each tanker belongs to one of four groups – GL, GH, BL, or BH. Each tanker $j$ knows its group and the sanctioning authority's resource allocation mechanism as in Eq.(8), but does not know $a_j$ (i.e., exactly how intensely the authority is monitoring it). Given Eq.(11), we can obtain the expected sanction probabilities for tankers in each group, conditional on their information set. Adding to the group index a 0 or 1, depending on whether the tanker is currently transporting sanctioned oil (i.e., whether

$\mathbb{I}$(deal with R) is 0 or 1), these conditional probabilities are:

$$\begin{aligned}
\bar{w}_{GL0} &= 0, & \bar{w}_{GL1} &= \mathbb{E}[a|G,L]\cdot\pi_2,\\
\bar{w}_{GH0} &= 0, & \bar{w}_{GH1} &= \mathbb{E}[a|G,H]\cdot\pi_2,\\
\bar{w}_{BL0} &= \mathbb{E}[a|B,L]\cdot\pi_1, & \bar{w}_{BL1} &= \mathbb{E}[a|B,L]\cdot[\pi_1+\pi_2]\\
\bar{w}_{BH0} &= \mathbb{E}[a|B,H]\cdot\pi_1, & \bar{w}_{BH1} &= \mathbb{E}[a|B,H]\cdot[\pi_1+\pi_2],
\end{aligned} \tag{12}$$

where we again omit the tanker index $j$ for brevity. Note that $\bar{w}_{GL0} = \bar{w}_{GH0} = 0$ because Good tankers know that they will not be sanctioned due to past violations; this assumes that these tankers are always "cleared" if investigated for past violations.

**Clean exporters**: They calculate the expected probabilities of L and H tankers being sanctioned as follows:

$$\begin{aligned}
\bar{w}^C_L &= \frac{Q_{GL}}{Q_{GL}+Q_{BL}}\bar{w}_{GL0} + \frac{Q_{BL}}{Q_{GL}+Q_{BL}}\bar{w}_{BL0},\\
\bar{w}^C_H &= \frac{Q_{GH}}{Q_{GH}+Q_{BH}}\bar{w}_{GH0} + \frac{Q_{BH}}{Q_{GH}+Q_{BH}}\bar{w}_{BH0},
\end{aligned} \tag{13}$$

where $Q_{GL}$, $Q_{GH}$, $Q_{BL}$, and $Q_{BH}$ are the number of tankers in each group that Clean exporters hire in equilibrium. These probabilities imply that Clean exporters realize that the proportion of Good vs. Bad tankers among the tankers *that they hire* is likely to be different from the proportion of Good vs. Bad tankers among *all* tankers, as tankers of specific types could be more likely to engage with them.[8]

**Rogue exporters**: They only know tanker types but not their risk labels. So they calculate the expected sanction probabilities for the G and B tankers that they hire as:

$$\begin{aligned}
\bar{w}^R_G &= \frac{Q^R_{GL}}{Q^R_{GL}+Q^R_{GH}}\bar{w}_{GL1} + \frac{Q^R_{GH}}{Q^R_{GL}+Q^R_{GH}}\bar{w}_{GH1},\\
\bar{w}^R_B &= \frac{Q^R_{BL}}{Q^R_{BL}+Q^R_{BH}}\bar{w}_{BL1} + \frac{Q^R_{BH}}{Q^R_{BL}+Q^R_{BH}}\bar{w}_{BH1}.
\end{aligned} \tag{14}$$

[8]Note that here each party's beliefs depend on quantities that they do not know ex-ante (e.g., Clean exporters never observe $Q_{GL}$ and $Q_{BL}$ separately – they only observe the sum, as they cannot distinguish between G and B tankers). In a rational expectations equilibrium, however, each party will make guesses about these unknown components such that their beliefs are consistent with model solutions (and therefore with each other).

### 3.1.6 Exporters' optimization

The objective of the exporters is to maximize profits. The per-unit oil trade revenue for Clean exporters, $\tilde{r}$, and for Rogue exporters, $\tilde{r}^R$, are

$$\tilde{r} = r + \tilde{\epsilon}, \qquad \tilde{r}^R = r^R + \tilde{\epsilon}, \tag{15}$$

where $r$ and $r^R$ are mean values and $\tilde{\epsilon}$ is the aggregate shock, uniformly distributed over $[-\sigma, \sigma]$. For simplicity, we assume a fixed ratio $\frac{r^R}{r}$, which implies a constant discount on sanctioned oil.[9]

Exporters' costs include (i) information acquisition cost (only for Clean exporters), (ii) payments to tankers, and (iii) penalties if the tanker they employ is sanctioned. The sanction penalties are $z^C$ and $z^R$ for Rogue and Clean exporters, respectively. We assume that $z^R > z^C$, as a Rogue exporter's entire oil cargo could be impounded if a tanker carrying it is caught and sanctioned, which is significantly more costly than the delays and penalties facing a Clean exporter for hiring a Bad tanker that now gets sanctioned but is currently carrying unsanctioned oil.

**Clean exporters**: They first choose their detection technology (i.e., $\sigma_\xi$ and $\chi$), and then set prices $p_L$ and $p_H$. Given the i.i.d. environment, the decisions on detection technology and prices are taken period by period.

We first consider the Clean exporters' price decision. Since tankers and exporters are randomly matched in a pair (Section 3.1.2), each Clean exporter occupies a $\frac{1}{N}$ portion of the market, and does not compete with other Clean exporters. For a given detection technology, each Clean exporter optimally chooses prices to maximize profits:

$$\tilde{\Pi} = \max_{\{p_L, p_H\}} \frac{1}{N}(Q_{GL} + Q_{BL})(\tilde{r} - p_L - \bar{w}_L^C z^C) + \frac{1}{N}(Q_{GH} + Q_{BH})(\tilde{r} - p_H - \bar{w}_H^C z^C), \tag{16}$$

where $\bar{w}_L^C$ and $\bar{w}_H^C$ are expected sanction probabilities defined in Eq.(13). The $Q$'s are functions of the prices, which are derived from tankers' optimization (as described below).

We use backward induction to derive Clean exporters' detection technology. Let $\Pi$ denote the expected equilibrium profit, given these exporters' belief about Windward's precision $\tilde{\sigma}_v^{-2}$ (again, symmetry implies same equilibrium beliefs):

$$\Pi(\sigma_\xi, \chi; \tilde{\sigma}_v^2) = \mathbb{E}[\tilde{\Pi} | \tilde{\sigma}_v^2] \tag{17}$$

[9] Consistent with the data in the months surrounding the Refinitiv disclosure (see, e.g., Inside Shipping, 2024).

The optimization problem is then

$$\max_{\{\sigma_\xi,\chi\in\{0,1\}\}} \Pi(\sigma_\xi,\chi;\tilde{\sigma}_v^2)-\Omega(\sigma_\xi)-\chi\cdot P^W, \tag{18}$$

with optimality conditions

$$\frac{\partial\Pi(\sigma_\xi,\chi)}{\partial\sigma_\xi}=\frac{\partial\Omega(\sigma_\xi)}{\partial\sigma_\xi},\quad \chi=\begin{cases}0, & \text{if } \Pi(\sigma_\xi,1)-\Pi(\sigma_\xi,0)<P^W,\\ 1, & \text{if } \Pi(\sigma_\xi,1)-\Pi(\sigma_\xi,0)>P^W,\\ 0 \text{ or } 1, & \text{if } \Pi(\sigma_\xi,1)-\Pi(\sigma_\xi,0)=P^W.\end{cases} \tag{19}$$

**Rogue exporters**: They solve the optimization problem only for prices:

$$\max_{\{p_G^R,p_B^R\}} (Q_{GL}^R+Q_{GH}^R)(\tilde{r}^R-p_G^R-\bar{w}_G^R z^R)+(Q_{BL}^R+Q_{BH}^R)(\tilde{r}^R-p_B^R-\bar{w}_B^R z^R), \tag{20}$$

where $\bar{w}_G^R$ and $\bar{w}_B^R$ are expected sanction probabilities defined in Eq.(14).

### 3.1.7 Tankers' optimization

Tankers observe freight rates (fixtures) quoted to them by Clean and Rogue exporters and decide whom to engage with. Because a Good tanker becomes irreversibly Bad if it engages with Rogue exporters, tankers must consider their continuation values when making decisions.

Let $V_G$ and $V_B$ denote the values of Good and Bad tankers, and $V_{GL}$, $V_{GH}$, $V_{BL}$, $V_{BH}$ denote the values further conditional on risk labels. Let $\theta_B=\mathbb{P}[H|B]$ and $\theta_G=\mathbb{P}[H|G]$. The tanker values satisfy the following relations:

$$V_G=\theta_G V_{GH}+(1-\theta_G)V_{GL},\quad V_B=\theta_B V_{BH}+(1-\theta_B)V_{BL}. \tag{21}$$

The tankers' Bellman equations are

$$\begin{aligned}
V_{GL}&=\mathbb{E}\Big[\max\Big\{p_L-\bar{w}_{GL0}\cdot z+\beta V_G,\quad p_G^R-\tilde{c}-\bar{w}_{GL1}\cdot z+\beta V_B\Big\}\Big],\\
V_{GH}&=\mathbb{E}\Big[\max\Big\{p_H-\bar{w}_{GH0}\cdot z+\beta V_G,\quad p_G^R-\tilde{c}-\bar{w}_{GH1}\cdot z+\beta V_B\Big\}\Big],\\
V_{BL}&=\mathbb{E}\Big[\max\Big\{p_L-\bar{w}_{BL0}\cdot z+\beta V_B,\quad p_B^R-\tilde{c}-\bar{w}_{BL1}\cdot z+\beta V_B\Big\}\Big],\\
V_{BH}&=\mathbb{E}\Big[\max\Big\{p_H-\bar{w}_{BH0}\cdot z+\beta V_B,\quad p_B^R-\tilde{c}-\bar{w}_{BH1}\cdot z+\beta V_B\Big\}\Big].
\end{aligned} \tag{22}$$

where $\bar{w}_{GL0}$, $\bar{w}_{GL1}$, etc., are expected sanction probabilities defined in Eq.(12), and $\tilde{c}$ is a random cost that a tanker incurs each period when dealing with Rogue exporters. Observed only by the tanker before choosing an exporter to deal with, this cost is assumed to be uniformly i.i.d. over $[0,\bar{c}]$ and allows us to obtain interior solutions for tanker supply. Such

a cost may arise from efforts to avoid detection, reflecting tanker-specific variables like its location, route, and violation technology.

The penalty imposed on a Bad tanker when sanctioned is $z$, assumed for simplicity to equal the tanker's entire value. Given discounting, this penalty is

$$z = \beta V_B. \tag{23}$$

The critical values $\tilde{c}_j$ at which tanker $j$ would be indifferent between dealing with a Clean or Rogue exporter are then:

$$c_{GL} = p_G^R - p_L - (\bar{w}_{GL1} - \bar{w}_{GL0})z - \beta(V_G - V_B), \qquad c_{BL} = p_B^R - p_L - (\bar{w}_{BL1} - \bar{w}_{BL0})z,$$
$$c_{GH} = p_G^R - p_H - (\bar{w}_{GH1} - \bar{w}_{GH0})z - \beta(V_G - V_B), \qquad c_{BH} = p_B^R - p_H - (\bar{w}_{BH1} - \bar{w}_{BH0})z.$$

For $\tilde{c}_j$ below (above) such a critical value, the tanker prefers to deal with a Rogue (Clean) exporter. Using these critical values, we derive that

$$V_G = \frac{\theta_G \mathbb{E}\left[\frac{(c_{GH})^2}{2\bar{c}} + p_H\right] + (1-\theta_G)\mathbb{E}\left[\frac{(c_{GL})^2}{2\bar{c}} + p_L\right] - [\theta_G w_{GH0} + (1-\theta_G)w_{GL0}]z}{1-\beta}, \tag{24}$$

$$V_B = \frac{\theta_B \mathbb{E}\left[\frac{(c_{BH})^2}{2\bar{c}} + p_H\right] + (1-\theta_B)\mathbb{E}\left[\frac{(c_{BL})^2}{2\bar{c}} + p_L\right] - [\theta_B w_{BH0} + (1-\theta_B)w_{BL0}]z}{1-\beta}, \tag{25}$$

where the expectation is over next period's aggregate shock $\tilde{\epsilon}$.

### 3.1.8 Equilibrium

A stationary equilibrium is characterized by: (i) Clean exporters' choice of detection technology (i.e., $\sigma_\xi$ and $\chi$); (ii) Clean and Rogue exporters' freight rates $p_G^R(\tilde{\epsilon})$, $p_B^R(\tilde{\epsilon})$, $p_L(\tilde{\epsilon})$, $p_H(\tilde{\epsilon})$, where $\tilde{\epsilon}$ refers to the aggregate shock in the current period; and (iii) tankers' decision rule based on the critical values $c_{GL}(\tilde{\epsilon})$, $c_{GH}(\tilde{\epsilon})$, $c_{BL}(\tilde{\epsilon})$, $c_{BH}(\tilde{\epsilon})$, which satisfy

- Optimality: Clean exporters optimally choose detection technology; Clean and Rogue exporters optimally set prices to maximize their profits; tankers maximize their values
- Market clearing: the market clears for all. In particular, for Clean exporters ($Q_{xy}$) and Rogue exporters ($Q_{xy}^R$) this implies, respectively

$$Q_{xy} = A_{xy}\left(1 - \frac{c_{xy}}{\bar{c}}\right), \quad Q_{xy}^R = A_{xy}\left(\frac{c_{xy}}{\bar{c}}\right), \quad xy \in \{GL, GH, BL, BH\},$$

where $A_{GL} = (1-\lambda)(1-\theta_G)$, $A_{GH} = (1-\lambda)\theta_G$, $A_{BL} = \lambda(1-\theta_B)$, and $A_{BH} = \lambda\theta_B$.

The model solution is described in detail in Internet Appendix F.

## 3.2 Model calibration and fit

### 3.2.1 Directly calibrated parameters

We focus on changes in market equilibrium in response to the disclosure. We directly calibrate seven parameters. The discount factor $\beta$ is set to 0.85, accounting for both time discounting and tanker depreciation.[10] Based on Lloyd's List (2023), we set the proportion $\lambda$ of Bad tankers to 0.1. We normalize $r$ to one – this is Clean exporters' mean oil trade revenue per voyage, or, more precisely, their profits before shipping costs. This revenue amounts to 15.5 mln for an average tanker (i.e., 0.8 million barrels at \$70 per barrel with a 27.7% before-shipping profit margin – this was the oil industry's operating margin for 2023 Q2 as per CSI Market, 2023). The mean revenue of Rogue exporters is $r^R = (70 \times 0.277 - 4)/(70 \times 0.277) = 0.794$, assuming a \$4 per barrel discount (as for Russian exports to India in the second half of 2023, Reuters, Sept. 2023).

Given that OFAC's civil monetary penalties typically depend on the amount of the sanctions-violating transaction (CFR, Appendix A to Part 501), we set the sanction penalty $z^C$ for Clean exporters dealing with Bad tankers equal to the freight cost they pay, i.e., around \$3 mln in 2023 (e.g., Inside Shipping, 2024, 0.8 million barrels at \$4/barrel), which implies $z = 3/15.5 = 0.194$. We set the sanction penalty $z^R$ for Rogue exporters based on the median value from actual cases of seized oil (e.g., U.S. Attorney's Office, 2024), resulting in $z^R = 50/15.5 = 3.226$.

### 3.2.2 Parameters derived from moments matching

We calibrate the remaining nine parameters by matching nine moments. These parameters are: the per-unit-resource sanction probabilities $\pi_1$ and $\pi_2$, max operation cost when dealing with Rogue exporters $\bar{c}$, volatility of oil trade revenue $\sigma$, precision of the public signal $\sigma_\zeta^{-2}$, precision of the authority's signal $\sigma_u^{-2}$, precision of the Windward's signal $\sigma_v^{-2}$, Clean exporters' belief on Windward's signal $\tilde{\sigma}_v$, and coefficient on information acquisition cost $\kappa$.

To identify the sanction probabilities, $\pi_1$ and $\pi_2$, and precision of the authority's signal, $\sigma_u^{-2}$, we mainly rely on three moments: actual proportion of sanctioned tankers, the disclosure's effect on tankers' fixtures, and the disclosure's effect on tankers' routes (these come from the DiD estimations in Table 2). These moments are particularly relevant because the actual proportion of sanctioned tankers mainly depends on $\pi_1$ and $\pi_2$, and the disclosure effects are determined by the authority's signal precision (as the disclosure releases a noisy version of the authority's signal) and the sanction probabilities $\pi_1$ and $\pi_2$.

[10] Assuming a time-discounting rate of 0.95, a 10% scrap value, and a 20-year further life cycle (given these are 10-15 year old tankers) gives $0.95 \times \exp(\log(0.1)/20) \approx 0.85$.

We identify the maximum operation cost $\bar{c}$ and the volatility of oil trade revenue $\sigma$ mainly by matching two moments: the volatility-to-mean ratios for low- and high-risk tanker fixtures. The two moments facilitate identification because $\bar{c}$ mainly contributes to the levels of low- and high-risk tanker fixtures, while $\sigma$ contributes to the volatility of fixtures.

We use the two pseudo-$R^2$s of sanction predictability, based on the pure-public signal and the Windward-and-public-data combined signal, to pin down the precision of the public signal, $\sigma_\zeta^{-2}$, and the Windward signal, $\sigma_v^{-2}$. We use Windward's market value to uncover Clean exporters' belief in Windward's precision $\tilde{\sigma}_v^{-2}$, as explained in the following section.

Finally, we identify the coefficient on information acquisition cost $\kappa$ by assuming this cost to be the major component of compliance costs and then match the ratio of non-compliance costs to compliance costs (see Secureframe (2025) and Globalscape). The moments and the parameters derived from moments matching are in Panels B and C of Table 3.

### 3.2.3 Calibrating $\tilde{\sigma}_v$

Recall that in Section 3.1.3 we left flexible the Clean exporters' beliefs about Windward's precision $\sigma_v^{-2}$. The most straightforward assumption would be to impose full rational expectations, i.e., $\tilde{\sigma}_v = \sigma_v$.

However, adopting this assumption results in a significant divergence between the model-implied values and the observed data, most notably with respect to Windward's valuation. Specifically, under this assumption our model would imply the valuation of its signal about suspect tankers to be \$5,405 mln in July 2023, whereas in fact the total equity valuation of Windward at that time was \$42 mln.

This divergence is so substantial that it is robust to many alternative modeling assumptions about the structure of the shipping market. Intuitively, a data provider like Windward's valuation depends on the value of the signal they provide, which, in turn, depends on market participants' expected incremental profits from buying that signal, which is at most \$27 mln in July 2023 given stock market prices.[11] This suggests that no market participant believed that buying Windward's signal would increase their profits beyond this amount (i.e., a high $\tilde{\sigma}_v$). However, as demonstrated in Section 2.2, the revenue of high-risk tankers declined by 13% following the Refinitiv disclosure, and these tankers started avoiding sanctioned exporters. Both of these findings indicate a large spillover of high-risk tankers in the Clean oil shipping market, reducing Clean exporters' shipping costs. Such a reduction, in turn, would imply that their incremental profits could be significantly higher than \$27 mln. But had

[11] While Windward's total equity valuation was \$42 mln, it had other lines of business besides the suspect tanker list. In the Internet Appendix F.4 we show how we can derive an upper bound of the valuation specific to its suspect tanker list.

these exporters anticipated such large incremental profits (i.e., had they correctly believed $\tilde{\sigma}_v = \sigma_v$), they should have been willing to pay much more for Windward's data, and its valuation should have accordingly been much higher.

Moving away from the equality $\tilde{\sigma}_v = \sigma_v$, however, brings forth the challenge of pinning down what exact value to assign to $\tilde{\sigma}_v$ in the calibration among the many possibilities. Next we explain how we do this, by referring back to Windward's equity valuation.

Clean exporter $i$'s willingness-to-pay for Windward's signal, $\tilde{\delta}_i$, depends on her belief about Windward's precision $\tilde{\sigma}_v^{-2}$ as follows:

$$\tilde{\delta}_i(\tilde{\sigma}_v^2) = \Pi(\sigma_{\xi,i}, 1; \tilde{\sigma}_v^2) - \Pi(\sigma_{\xi,i}, 0; \tilde{\sigma}_v^2), \tag{26}$$

where $\Pi(\cdot; \tilde{\sigma}_v^2)$ is the expected profit as in Eq.(18) and its argument one or zero denotes whether Windward's signal is bought or not. Let the true value of Windward's information be $\delta$. We assume that Clean exporters' prior belief about $\delta$ is distributed $\mathbb{N}(\mu_\delta, \Sigma_\delta)$. Additionally, each Clean exporter receives a noisy signal (e.g., from tankers newly sanctioned in every period or from Windward providing a trial version of their list before purchase) about the true value of $\delta$: $x_i = \delta + \eta_i$, with independent noise $\eta_i \sim \mathbb{N}(0, \Sigma_\eta)$. Each Clean exporter $i$ rationally updates her willingness-to-pay based on her signal $x_i$ and Windward's market value $M^W$:

$$\tilde{\delta}_i = \mathbb{E}[\delta | x_i, M^W]. \tag{27}$$

The absence of near-arbitrage opportunities implies that

$$M^W = \sum_{i=1}^{N} \tilde{\delta}_i + \epsilon^W, \tag{28}$$

where $\epsilon^W \sim \mathbb{N}(0, \Sigma_\epsilon)$ comes from noise traders.[12] Solving Eq.(27) and Eq.(28) gives

$$\tilde{\delta}_i = \mu_\delta + A(x_i - \mu_\delta) + B\left(\frac{M^W}{N} - \mu_\delta\right), \tag{29}$$

where $A$ solves $\frac{\Sigma_\eta}{\Sigma_\epsilon}(N-1)(N + \frac{\Sigma_\eta}{\Sigma_\delta})A^3 + A - 1 = 0$ and $B = \frac{N(N-1)\Sigma_\eta A^2}{\Sigma_\epsilon + N(N-1)\Sigma_\eta A^2}$. The derivation details are in the Internet Appendix F.3.

Note that, since Windward is a start-up with very short history and sanction rules are constantly updated yet little new information on violators is provided every period, a Clean exporter who has not yet decided to buy Windward's signal (and maybe has a trial version with a small subset of the data to decide whether to buy) would likely face an information environment characterized by (i) sufficiently noisy priors on Windward's accuracy ($\Sigma_\delta/\Sigma_\eta \to$

[12] For example, if $M^W << \sum_{i=1}^{N} \tilde{\delta}_i$, all Clean exporters can buy Windward together to realize the surplus.

$\underline{c} > 0$), and (ii) a sufficiently noisy signal $x_i$ ($\Sigma_\epsilon/\Sigma_\eta \to 0$). In the Internet Appendix F.3, we show that under these circumstances $A \to 0$ and $B \to 1$, and hence

$$\tilde{\delta}_i \to \frac{M^W}{N}. \tag{30}$$

Eq.(30) combined with $\tilde{\delta}_i(\tilde{\sigma}_v^2) = \Pi(\sigma_\xi, 1; \tilde{\sigma}_v^2) - \Pi(\sigma_\xi, 0; \tilde{\sigma}_v^2)$ allows us to uncover Clean exporters' belief from Windward's market value. In such an equilibrium, Clean exporters are indifferent between $\chi = 0$ and $\chi = 1$ (i.e., buying or not buying Windward's signal), as per Eq.(19), assuming that Windward is competitively priced (i.e., it's value reflects the entire surplus it generates).

### 3.2.4 Implications of $\tilde{\sigma}_v \neq \sigma_v$

The fact that Clean exporters' ex-ante $\tilde{\sigma}_v$ did not equal the true $\sigma_v$ reveals an important friction in the market. Had these exporters been able to recognize Windward's true accuracy in detecting violators, they would have expected much higher incremental profits from its information, and hence would have been willing to pay for Windward accordingly. A significantly higher valuation would have motivated the creation of similar companies who could charge appropriately for detecting suspected violators in different domains. This would represent an endogenous market-based resolution of the detection challenges.

However, as we demonstrated above, this solution in not supported by the data. With diffuse priors about Windward's precision and inadequate additional information to update beliefs over time, it is rational for Clean exporters to try to infer this precision by learning from the company's market value. And this means that a wide range of beliefs and associated valuations can be justified in equilibrium. If the market somehow underestimated Windward's value, noisy learning would have prevented convergence to the truth. Such a situation, where Windward gets stuck in a value trap, could then justify the company releasing their signal at very low cost to demonstrate its true accuracy (and thus benefit its other lines of business). In fact, Windward's stock market valuation had indeed dropped to half of its IPO value and stayed at that level or below for over a year before the disclosure – i.e., the undervaluation was indeed persistent.

Two further features in the data are consistent with this interpretation: (1) Windward tried to maximize the reach of its disclosure by releasing the list via Refinitiv Eikon, a widely used platform, rather than on its own website. This is consistent with the point of the disclosure being publicity. (2) The demonstration led to increased demand for Windward's other services, as the market revised its beliefs about the firm's overall detection capabilities, helping offset any revenue lost from giving away the sanctions list for free. Following the

disclosure, Windward's valuation actually rose sharply, and the firm was ultimately taken private in March 2025 at a valuation of $270 mln. This indicates that Windward was right about the market updating its beliefs positively about their business in such a way that the disclosure would end up being value increasing, i.e, ex-ante the disclosure was justified given the later value increase.

### 3.2.5 Model fit

We report the moment matching results in Table 3 Panel B. In particular, the disclosure effects on high-risk tankers' fixture and route in the model (i.e., -0.126 and -0.184) are close to the values in the data (-0.13 and -0.17), and the remaining moments are similarly well matched.

We also compare the values of other model variables that are not targeted in our calibration with the data. First, the volatility of oil trade revenue derived from moments matching is $\sigma$ = 0.26, consistent with the Oil Volatility Index OVX around 28% in 2023 (OVX measures oil price volatility, and hence, earnings volatility). Second, the model-implied freight rate premium paid by Rogue exporters relative to Clean exporters is around between 100-200%, as per Panel D in Table 3. This is closer to the high end of the sanctions premium reported in the media for Russian oil delivered to Indian and Chinese ports in August 2023 (e.g., Inside Shipping, 2024, which reports calculations from Argus Media). Third, model-based tanker values range between $40 to $60 mln, comparable to the actual values for 15-year-old tankers as per Xclusiv (2023), and noting that the average age of the crude oil tanker fleet in 2023-2024 was 13-14 years, per AXSmarine (2025).

Overall, our model-based estimates appear to align reasonably well with their corresponding values reported by alternative sources, offering some confidence in the reliability of the quantities generated by the model.

## 3.3 Model implications

In Figure 4, we present the model implications of the Refinitiv disclosure by varying the precision of Windward's signal, $\sigma_v^{-2}$. We focus on the pre-period equilibrium where almost no Clean exporters buy Windward's signal – which we model as Windward's signal precision being zero before disclosure – and the post-period where everyone has access for free. The two vertical dashed lines indicate these two equilibria, respectively.

To better interpret the disclosure effect, we quantify how the disclosure changes Clean exporters' detection accuracy, measured by the pseudo-$R^2$ for predicting Bad tankers. Our model can provide this calculation for two reasons. First, we can empirically estimate the

Table 3: **Model parameters, moments matching, and implied quantities**

This table presents the directly-calibrated model parameters in Panel A, the moments that we match in Panel B, the parameters derived from moments matching in Panel C, the imputed quantities in Panel D. These results are discussed in Sections 3.2 and 3.3.

Panel A: Directly calibrated parameters

| Variable | Calibration method | Symbol | Value |
|---|---|---|---|
| Discount factor | time-discounting (0.95) × depreciation | $\beta$ | 0.85 |
| Proportion of Bad tankers | see Section 3.2.1 | $\lambda$ | 0.1 |
| Mean oil trade revenue (Clean exporters) | \$15.5m, normalized to one | $r$ | 1 |
| Mean oil trade revenue (Rogue exporters) | \$4 per barrel discount | $r^R$ | 0.794 |
| Sanction penalty (Clean exporters) | \$3m, comparable to shipping costs | $z^C$ | 0.194 |
| Sanction penalty (Rogue exporters) | \$50m, based on the median value from actual cases of seized oil | $z^R$ | 3.226 |

Panel B: Matched moments

| Moment | Model analog | Data | Model |
|---|---|---|---|
| Proportion of sanctioned tankers (average of pre- and post-Windward periods) | $\frac{\mathbb{E}[Q^{pre}_{Sanc}+Q^{post}_{Sanc}]}{2}$ | 0.7% | 0.89% |
| Disclosure's effect on high-risk tankers' fixtures (DiD estimation in Table 3) | $\frac{\mathbb{E}[p^{post}-p^{pre}\|\text{Post H}]}{\mathbb{E}[p^{pre}\|\text{Post H}]}$ | -0.13 | -0.126 |
| Disclosure's effect on high-risk tankers' routes (DiD estimation in Table 4) | $\frac{\mathbb{E}[b^{post}-b^{pre}\|\text{Post H}]}{\mathbb{E}[b^{pre}\|\text{Post H}]}$ | -0.17 | -0.184 |
| Annualized volatility of low-risk tankers' fixtures before the disclosure (divided by mean) | $\frac{\text{std}[p^{pre}_L]}{\mathbb{E}[p^{pre}_L]}$ | 0.77 | 0.788 |
| Annualized volatility of high-risk tankers' fixtures before the disclosure (divided by mean) | $\frac{\text{std}[p^{pre}_H]}{\mathbb{E}[p^{pre}_H]}$ | 0.81 | 0.829 |
| Pseudo-$R^2$ for predicting to-be-sanctioned tankers with the public signal | See Internet Appendix F | 4.3% | 4.4% |
| Pseudo-$R^2$ for predicting to-be-sanctioned tankers with the public-Windward-combined signal | See Internet Appendix F | 15.4% | 15.4% |
| Valuation of the Windward's signal (based on the share price one day before the disclosure) | $\Pi(\sigma^*_\xi,1)-\Pi(\sigma^*_\xi,0)$ | \$27m | \$27m |
| Non-compliance costs divided by compliance costs (see Secureframe (2025)) | $\frac{\mathbb{E}[(Q^{pre}_{GL}+Q^{pre}_{BL})\bar{w}^C_L z^C+(Q^{pre}_{GH}+Q^{pre}_{BH})\bar{w}^C_H z^C]}{\Omega(\sigma^*_\xi)}$ | 2.71 | 2.70 |

Panel C: Parameters derived from moments matching

| Variable | Symbol | Value |
|---|---|---|
| Sanction probability per unit of resource (previous violation) | $\pi_1$ | 2.6% |
| Sanction probability per unit of resource (current violation) | $\pi_2$ | 3.0% |
| Max operation cost when dealing with Rogue exporters | $\bar{c}$ | 0.89 |
| Volatility of oil trade revenue | $\sigma$ | 0.26 |
| Noise volatility in the public signal | $\sigma_\zeta$ | 0.68 |
| Noise volatility in the authority's signal | $\sigma_u$ | 0.52 |
| Noise volatility in the Windward's signal (true value) | $\sigma_v$ | 0.10 |
| Noise volatility in the Windward's signal (Clean exporters' ex-ante belief) | $\bar{\sigma}_v$ | 8.28 |
| Coefficient on information acquisition cost ($\times 10^3$) | $\kappa$ | 2.30 |

Panel D: Imputed quantities (price in \$ mln per trip)

| Period | $\mathbb{E}(p_H)$ | $\mathbb{E}(p_L)$ | $\mathbb{E}(p^R_B)$ | $\mathbb{E}(p^R_G)$ | $V_B$ | $V_G$ |
|---|---|---|---|---|---|---|
| Pre-Windward | 2.85 | 2.99 | 4.76 | 9.60 | 43.90 | 61.14 |
| Post-Windward | 2.59 | 2.91 | 4.60 | 9.74 | 40.48 | 59.08 |

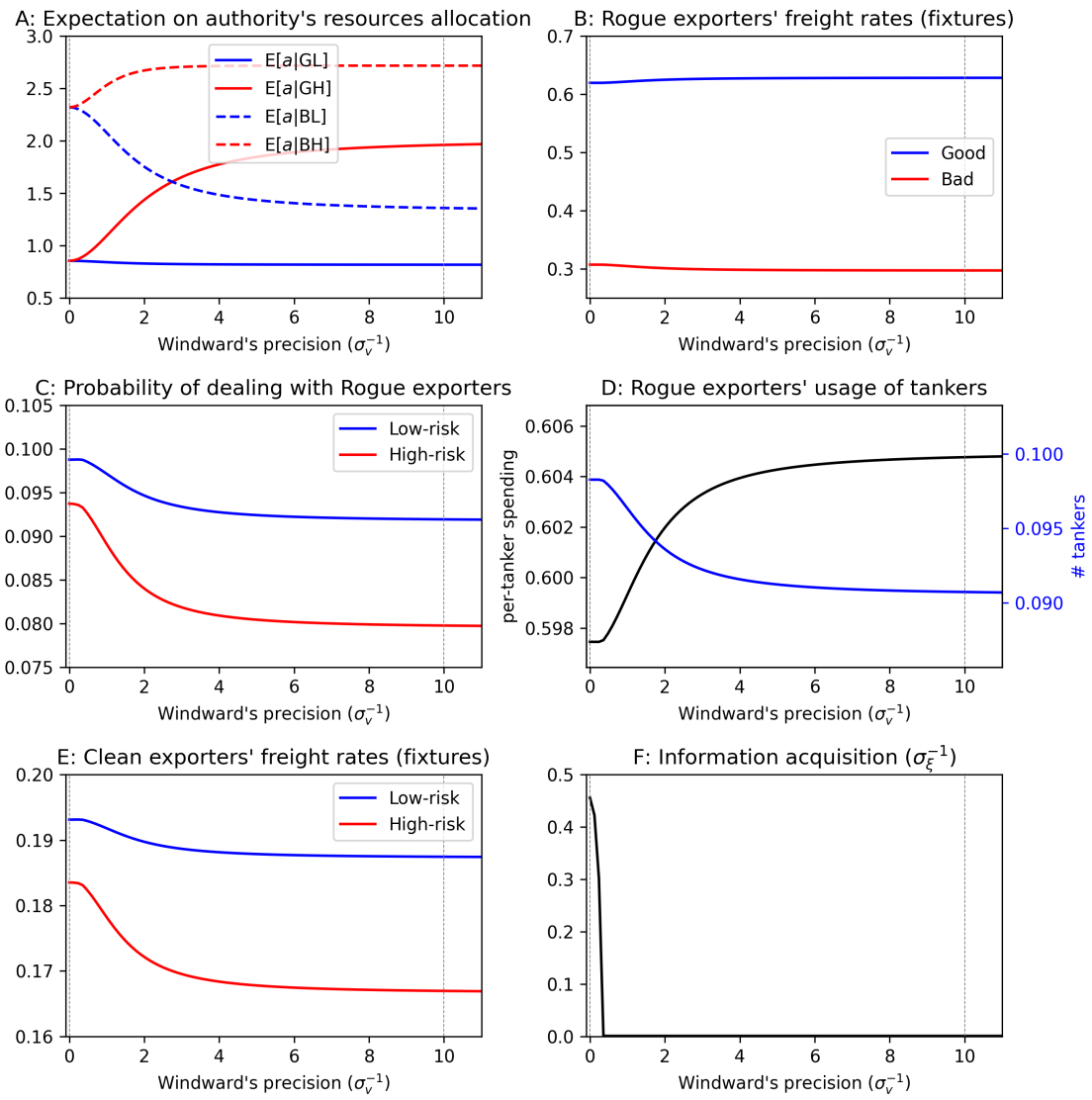



Figure 4: This figure presents the predictions from the calibrated model from Section 3. We solve the model by discretizing the aggregate shock $\tilde{\epsilon}$ on a grid. We plot the average model outcomes for the values of $\tilde{\epsilon}$ as functions of pseudo-$R^2$, which is backed out from the classification precision $\rho$ in the model. In each subplot, two vertical dashed lines indicate the pre- and post-Windward precisions, which correspond to 18% and 42% pseudo-$R^2$ of predicting Bad tankers (or 4.3% and 15.4% pseudo-$R^2$ of predicting to-be-sanctioned tankers), respectively.

pseudo-$R^2$ for predicting which tankers will be sanctioned – 4.3% in the pre-period and 15.4% in the post-period, as used in moments matching – which in our framework is jointly determined by Clean exporters' detection accuracy and the authority's signal precision. Second, under our information structure – where the public signal $s^{C,com}$ is independent of the authority's signal $s^A$, and Windward's signal $s^W$ is a noisy version of $s^A$ – our model can identify Clean exporters' detection accuracy separately from the authority's signal precision.

The assumed independence between $s^A$ and $s^{C,com}$ is consistent with the authority's encouragement of market participants' own due diligence, and the correlation between $s^A$ and $s^W$ is supported by Reuters (Mar. 2016), and Wired (Mar. 2020). The pseudo-$R^2$ formula is in Internet Appendix F. Our calibration indicates that the disclosure raises Clean exporters' pseudo-$R^2$ for predicting Bad tankers from 18% to 42%.

Panel A of Figure 4 illustrates tankers' expectations about the authority's resource allocation. First, because the authority allocates resources based on a noisy signal about tanker types, Bad tankers rationally anticipate that they tend to face more monitoring than Good tankers, for each risk label. Second, as the precision of Windward's signal improves, High-risk (Low-risk) tankers expect that they will face more (less) monitoring. This is because in the post-period, risk labels are derived based on Windward's signal, which contains information about the authority's signal. Therefore, with higher precision of Windward's signal tankers can learn more about the authority's resource allocation from risk labels. This learning is driving the disclosure effect on route changes.

Panel B shows that as Windward's precision improves, Rogue exporters pay slightly higher freight rates to hire Good tankers (i.e., $\mathbb{E}[p_G^R]$ increases) and lower rates for Bad tankers (i.e., $\mathbb{E}[p_B^R]$ decreases). This is because, with better information, Clean exporters are more likely to detect Bad tankers, and hence the bargaining power of these tankers relative to Rogue exporters decreases, leading to lower freight rates. Knowing that, Good tankers require higher compensation from Rogue exporters for dealing with them and thus turning into Bad tankers thereafter.

Panel C shows the probabilities of tankers engaging in transactions with Rogue exporters. High-risk tankers are less likely to engage with Rogue exporters. Importantly, as Windward's precision improves, the likelihood of engaging with Rogue exporters declines more sharply for High-risk tankers, as those tankers learn from their risk labels that the sanctioning authority now allocates more monitoring resources to them. These results align with our empirical findings on route changes.

Panel D sheds light on the aggregate implications for Rogue exporters. The black line shows that the per-tanker spending of Rogue exporters increases, which is due mostly to the increase in the freight rates paid by Rogue exporters to Good tankers (from Panel B in the figure). This increase is not offset by the drop in freight rates that these exporters pay to Bad tankers because there are many more Good tankers than Bad tankers.

The blue line in Panel D shows that as Windward's precision improves, the total number of oil tankers engaging with Rogue exporters decreases, consistent with Panel C. Quantitatively, the information shock accounts for a 7.7% drop in Rogue exporters' total hiring of tankers. This shortfall, of course, can translate into a drop in exports of sanctioned oil

itself, but the magnitude of that drop may not be one-for-one. This is because some of these tankers can be replaced by other means, e.g., through oil pipelines – although this likely involves higher costs of transportation (if the costs were lower, then Rogue exporters would have used these means rather than tankers in the pre-period).

Panel E in the figure shows the relation between Windward's precision and the equilibrium freight rates (i.e., the fixtures $\mathbb{E}[p_L]$ and $\mathbb{E}[p_H]$) from Clean exporters. First, as precision increases, High-risk tankers' freight rates sharply decrease. This is because with higher precision the H label indicates a higher probability of a Bad tanker, so Clean exporters require a larger discount for hiring such a tanker.

Additionally, Panel E reveals a counterintuitive result: as precision increases, Low-risk tankers' freight rates also decrease. This outcome arises from two opposing forces, which we derive from the first-order conditions of exporters' optimization. The first is an information effect: with improved precision, the L label more desirable as it indicates lower sanction risk. The second is a competition effect: as precision increases, High-risk tankers that are increasingly excluded from the Rogue exporters' market now relocate to the Clean market, intensifying competition and putting downward pressure on freight rates for Low-risk tankers. Our quantitative analysis indicates that the competition effect outweighs the information effect, resulting in a net decline in freight rates for Low-risk tankers. In the following section, we empirically test this counterintuitive prediction using a shift-share design.

Panel F in the figure shows that as Windward's precision improves, Clean exporters have lower incentive to acquire private information. Our quantitative analysis here shows that as Windward's precision increases, Clean exporters quickly switch their information source to Windward's signal. The substitution occurs because of the decreasing marginal benefits of information for Clean exporters. Intuitively, this follows from the fact that the generation of private information for violator detection is expensive in the data; and the fact that when Windward discloses its list, the market responds substantially – implying that the Clean exporters' private signal had not managed to generate Windward's information on its own. This suggests a high cost-to-accuracy ratio of the private signal; so, as long as Windward detection information is at least somewhat precise, Clean exporters stop generating their signal.

This intuition supports the view that providing better public information crowds out efforts to gather useful private information – which the sanctioning authority might itself lack.[13] However, given that this private information had a high cost-to-accuracy ratio, the

[13]Consider, for example, a company asked to deliver satellite surveillance equipment to the address of a pizza shop; the firm should reasonably suspect foul play (see Huneke (2023) for further examples) and flag that counter-party; this is something that the authority cannot do.

disclosure of Windward's more accurate signal improves enforcement overall (Section 3.3 contains further details).

Finally, in Table 4, we estimate the aggregate (annual) effect of the information shock on the four relevant parties in terms of dollar values. We analyze the changes in earnings for exporters and tankers between the pre- and post-Windward periods. We find that the shipping cost for Rogue exporters increases by 1.2% (from 9.26 to 9.37 mln per tanker-trip) due to the increase in $p_G^R$, while the cost for Clean exporters decreases by 3.5% (from 2.98 to 2.87 mln per tanker-trip) due to reductions in both $p_L$ and $p_H$. On the tanker side, fixture income decreases by 5.5% for Bad tankers (from 3.05 to 2.88 mln per tanker-trip) and by 2.0% for Good tankers (from 3.66 to 3.58 mln per tanker-trip). To estimate the aggregate annual effect, we do the following back-of-the-envelope calculation: we multiply the above dollar values by the number of tankers and 3 trips per year (from the data). Therefore, the disclosure leads to $1.1 bln annual losses for Rogue exporters and Bad tankers, and $1.9 bln overall gain for Clean exporters and Good tankers.

Table 4: **Model-implied overall effects**

This table reports the aggregate annual effect on earnings for the four relevant parties – Rogue exporters, Clean exporters, Bad tankers, and Good tankers (in mln dollars). We estimate by multiplying per-tanker-trip earnings by the number of tankers and 3 trips per year (the average number of trips per tanker from our fixtures data).

| | Sanctions violators | | Sanctions-compliant agents | |
|---|---|---|---|---|
| | R exporters (price paid) | Bad tankers (price received) | C exporters (price paid) | Good tankers (price received) |
| Pre-Windward (per tanker-trip) | 9.26 | 3.05 | 2.98 | 3.66 |
| Post-Windward (per tanker-trip) | 9.37 | 2.88 | 2.87 | 3.58 |
| Percentage change | 1.2% | -5.5% | -3.5% | -2.0% |
| Aggregate effect (annual, $ mln) (-/+ indicates loss/gain) | -562 | -510 | 4,299 | -2,443 |
| Total (annual, $ mln) (-/+ indicates loss/gain) | -1,072 | | 1,855 | |

## 3.4 Counterfactual analysis: Increasing penalties under noisy detection

Finally, to further explore implications for enforcement design, we use our model to study a key counterfactual: what would happen if the sanctioning authority increased the penalty

($z$) on Clean exporters who were found to be using sanctions-violating tankers.[14] This scenario aligns with much of the focus in political and media discussions, as well as literature suggesting that aligning the incentives of all parties with that of the sanctioning authority should enhance enforcement.

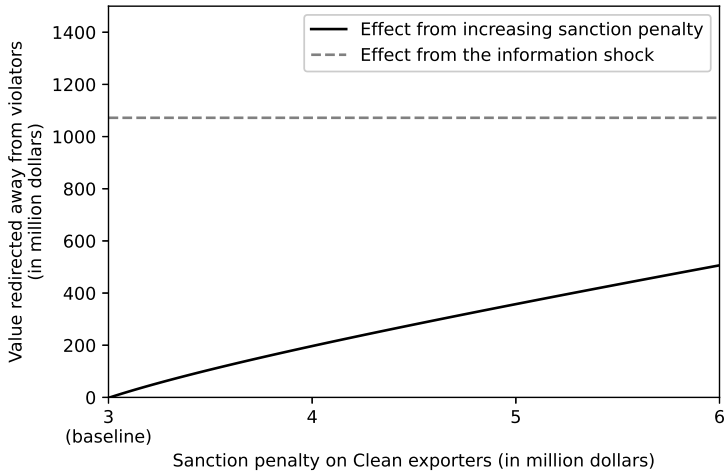



Figure 5: This figure presents the counterfactual analysis of increasing the sanction penalty. The $x$-axis is the sanction penalty for Clean exporters (i.e., parameter $z$) expressed in million dollars. The $y$-axis is the model-implied effects of the value redirected away from violators, as calculated in Table 4. For comparison, we plot the effect from the information shock using a dotted line.

In Figure 5, we conduct this analysis by varying $z$ from 0.258 (4 mln dollars) to 0.516 (8 mln dollars) and fixing detection precision at the pre-Windward level. We focus on the effect of the value redirected away from violators, as calculated in Table 4.

The figure reveals that increasing sanction penalties in the presence of noisy detection has little effect on enforcement. This is due to two reasons: first, as Clean exporters cannot perfectly observe tanker compliance status, increasing penalties raises the expected cost of inadvertently hiring a violating tanker. In response, Clean exporters reduce their overall demand for tanker services to manage this risk. This hurts both Good and Bad tankers. Of course, Bad tankers are hurt more, as Clean exporters start especially avoiding high-risk labeled ones that are more likely to be Bad. However, this shrinkage in demand from the compliant sector increases tanker supply and lowers cost for Rogue exporters, making the total value transferred away from violators in response to higher penalties negligible.

[14] Recall that for Rogue exporters, we already assumed that the sanctioning authority impounds the entire oil cargo – i.e., the penalty $z^R$ cannot be increased any further. Hence we focus on Clean exporters here.

## 3.5 Empirically testing the counter-intuitive model implication

In this final section, we test the main counter-intuitive model implication that after the disclosure, low-risk tankers receive lower fixtures from Clean exporters due to increased competition in that market (as shown in Panel E of Figure 4). We conduct tests exploiting charterer-level variations.

Ideally, we would want to measure the incremental tanker supply to each charterer after the information shock, but this quantity is unobservable. We circumvent this issue in designing our test by relying on the intuition of "shift-share instruments" (Bartik, 1991). Let $g_k$, for $k \in 1, \ldots, K$, be the varying exposures to increased competition (i.e., "shifts") of $K$ groups of tankers. Let also $s_{ik}$ be the proportion (i.e., "share") of group $k$ tankers that charterer $i$ deals with ($\sum_k s_{ik} = 1$). Then the share-weighted sum of exposures $\sum_k s_{ik} g_k$ serves as an ex-ante measure of charterer-level competition.[15]

To ensure variation in the exposures $g_k$, we group tankers by product type – i.e., refined products (e.g., gasoline and diesel), unrefined (e.g. crude oil), and unclassified products, hence $K = 3$. Among the refined-product tankers in Windward's list, 24% are high/moderate-risk tankers, while this proportion is 36% for unrefined, and 34% for unclassified products tankers. This variation suggests that charterers with more unrefined or unclassified products are more likely to experience increased supply from risky tankers after the disclosure. Note that this strategy nets out other common/aggregate changes in fixtures by exploiting charterer-level variation in exposures to supply shocks.

Using this measure, we focus on low-risk tankers and run the following regression:

$$fixture_{c,i,k,t} = \beta_0 \times \mathbb{I}_{\{\text{High-competition}\}} + \beta_1 \times \mathbb{I}_{\{\text{High-competition}\}} \times \mathbb{I}_{\{t \geq 0\}} + \text{FE} + \epsilon_{c,i,k,t}. \quad (31)$$

The dependent variable is low-risk tankers' fixtures, and $\mathbb{I}_{\{\text{High-competition}\}}$ equals one if the charterer-level shift-share competition measure exceeds a cross-sectional cutoff (e.g., top 50%, 40%, or 30%), and is zero otherwise. Our model predicts a negative DiD coefficient $\beta_1$.

Estimation results in Table 5 are consistent with model predictions: higher competition corresponds to lower freight rates for low-risk tankers after the disclosure. The results are robust across various cutoffs for the high-competition indicator, with estimated effects ranging from 12.9 to 15.9 WS units. These results not only provide an out-of-sample test for our model, but suggest that policymakers should be mindful of negative spillovers on compliant parties in equilibrium.

[15] For this measure to work, one would need stickiness in the charterer-tanker relationship. We verify this in the data – 40% of the tankers engage with charterers that they have worked with in the preceding two years.

Table 5: **Testing the counter-intuitive model implication: Competition spillovers**

We test the model's implication that the disclosure reduces low-risk tankers' freight rates (competition effects), exploiting charterer-level variation in a DiD framework. Since tankers carrying unrefined and unclassified products are more likely to be classified as risky by Windward, the charterers with a larger proportion of such tankers are expected to face relatively higher supply of low-risk tankers after the disclosure, leading to higher competition among these tankers and lower freight rates received by them. The dependent variable is low-risk tankers' fixtures. The explanatory variables are the high-competition indicator $\mathbb{I}_{\{\text{High-competition}\}}$ and its interaction with the post-period indicator. $\mathbb{I}_{\{\text{High-competition}\}}$ equals one if the charterer-level shift-share competition measure, as described in Section 3.5, exceeds certain cross-sectional cutoff (top 50, 40, or 30%), and zero otherwise. The standard errors are double clustered at the tanker and time×tanker-type levels. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

| | Dependent variable: Low-risk tankers' fixtures | | |
|---|---|---|---|
| | Cutoff = top 50% | Cutoff = top 40% | Cutoff = top 30% |
| $\mathbb{I}_{\{\text{High-competition}\}} \times \mathbb{I}_{\{t\geq 0\}}$ | -12.893** | -14.304** | -15.866** |
| | [-2.15] | [-2.48] | [-2.62] |
| Obs. (tanker-month) | 4,999 | 4,999 | 4,999 |
| Time × Tanker Type FE | Yes | Yes | Yes |
| Tanker FE | Yes | Yes | Yes |

# 4 Conclusion

This paper shows that enforcement failures primarily stem from third parties' inability to reliably identify violators, rather than from weak compliance incentives. We make three contributions.

First, we empirically show that better information on violators alone can improve enforcement outcomes. When a sanctions-risk list from the maritime analytics firm Windward.ai became public on Refinitiv Eikon, suspect tankers experienced sharp earnings declines and started avoiding sanctioned countries.

Second, we develop a dynamic structural model to quantify the market-wide consequences of improved detection. The model implies that $1.4 billion annually was redirected away from violators. Compliant exporters benefited, but — counterintuitively — compliant tankers saw lower earnings.

Third, we use the model to explain why, despite these large impacts, markets persistently undervalue information from vendors like Windward. The mechanism is a feedback loop between beliefs about detection accuracy and firm valuation.

These insights carry meaningful implications for policy design. If the core constraint on enforcement is information rather than incentives, then simply increasing penalties or tightening regulations will yield limited gains, as our model shows. Instead, improving

detection capabilities may be the most effective path toward strengthening the enforcement of market-based sanctions.

# References


Abadie, A. (2005), 'Semiparametric Difference-in-Differences Estimators', *Review of Economic Studies* **72**(1), 1–19.

Babina, T., Hilgenstock, B., Itskhoki, O., Mironov, M. and Ribakova, E. (2023), 'Assessing the impact of international sanctions on Russian oil exports', *Available at SSRN* .

Bai, X., Fernández-Villaverde, J., Li, Y., Xu, L. and Zanetti, F. (2025), 'The (Un)Intended Consequences of Oil Sanctions Through the Dark Shipping of Sanctioned Oil', *Working Paper* .

Baldwin, D. A. (2020), 'Economic statecraft: New edition', *Princeton University Press* .

Banerjee, S. (2011), 'Learning from prices and the dispersion in beliefs', *The Review of Financial Studies* **24**(9), 3025–3068.
**URL:** *https://doi.org/10.1093/rfs/hhr050*

Bartik, T. J. (1991), *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute for Employment Research.

Bond, P., Edmans, A. and Goldstein, I. (2012), 'The real effects of financial markets', *Annu. Rev. Financ. Econ.* **4**(1), 339–360.

Brancaccio, G., Kalouptsidi, M. and Papageorgiou, T. (2020), 'Geography, transportation, and endogenous trade costs', *Econometrica* **88**(2), 657–691.

Brancaccio, G., Kalouptsidi, M. and Papageorgiou, T. (2023), 'The impact of oil prices on world trade', *Review of International Economics* **31**(2), 444–463.

Brown, P. (2020), *Oil market effects from US economic sanctions: Iran, Russia, Venezuela*, Vol. 3, Congressional Research Service Washington, DC, USA.

Cipriani, M., Goldberg, L. S. and La Spada, G. (2023), 'Financial sanctions, swift, and the architecture of the international payment system', *Journal of Economic Perspectives* **37**(1), 31–52.

Cohen, L., Diether, K. and Malloy, C. (2013), 'Misvaluing innovation', *The Review of Financial Studies* **26**(3), 635–666.

Crozet, M., Hinz, J., Stammann, A. and Wanner, J. (2021), 'Worth the pain? Firms' exporting behaviour to countries under sanctions', *European Economic Review* **134**.

Dellavigna, S. and La Ferrara, E. (2010), 'Detecting illegal arms trade', *American Economic Journal: Economic Policy* **2**(4), 26–57.

Dimmock, S. G. and Gerken, W. C. (2012), 'Predicting fraud by investment managers', *Journal of Financial Economics* **105**(1), 153–173.

Dizaji, S. F. and van Bergeijk, P. A. G. (2013), 'Potential early phase success and ultimate failure of economic sanctions: A VAR approach with an application to Iran', *Journal of Peace Research* **50**, 721–736.

Early, B. R. and Preble, K. (2019), 'Enforcing Economic Sanctions: Analyzing How OFAC Punishes Violators of U.S. Sanctions', *Working Paper* .

Eaton, J. and Engers, M. (1992), 'Sanctions', *Journal of Political Economy* **100**, 899–928.

Eaton, J. and Engers, M. (1999), 'Sanctions: Some simple analytics', *American Economic Review* **89**, 409–414.

Edmans, A., Goldstein, I. and Jiang, W. (2012), 'The real effects of financial markets: The impact of prices on takeovers', *The Journal of Finance* **67**(3), 933–971.

Efing, M., Goldbach, S. and Nitsch, V. (2023), 'Freeze! Financial sanctions and bank responses', *Review of Financial Studies* **36**, 4417–4459.

Fisman, R. J., Marcolongo, G. and Wu, M. (2024), 'The undoing of economic sanctions: Evidence from the Russia-Ukraine conflict', *Working Paper* .

Fisman, R. and Wei, S.-J. (2004), 'Tax rates and tax evasion: evidence from "missing imports" in China', *Journal of Political Economy* **112**(2), 471–496.

Fisman, R. and Wei, S.-J. (2009), 'The smuggling of art, and the art of smuggling: Uncovering the illicit trade in cultural property and antiques', *American Economic Journal: Applied Economics* **1**(3), 82–96.

Goldstein, I. and Yang, L. (2019), 'Good disclosure, bad disclosure', *Journal of Financial Economics* **131**(1), 118–138.

Griffin, J. and Kruger, S. (2023), 'What is Forensic Finance?', *Available at SSRN* .

Griffin, J. M. and Maturana, G. (2016), 'Who facilitated misreporting in securitized loans?', *The Review of Financial Studies* **29**(2), 384–419.

Griffin, J. M. and Shams, A. (2018), 'Manipulation in the VIX?', *Review of Financial Studies* **31**(4), 1377–1417.

Haidar, J. I. (2017), 'Sanctions and export deflection: evidence from Iran', *Economic Policy* **32**, 319–355.

Hsieh, C.-T. and Moretti, E. (2006), 'Did Iraq cheat the United Nations? Underpricing, bribes, and the oil for food program', *Quarterly Journal of Economics* **121**, 1211–1248.

Hufbauer, G. C. and Jung, E. (2020), 'What's new in economic sanctions?', *European economic review* **130**, 103572.

Huneke, M. (2023), The Blind Men and the Elephant, Hughes Hubbard & Reed.

Huynh, L. D. T., Hoang, K. and Ongena, S. (2023), 'The impact of foreign sanctions on firm performance in Russia', *Swiss Finance Institute Research Paper* (23-115).

Itskhoki, O. and Mukhin, D. (2023), International sanctions and limits of Lerner symmetry, *in* 'AEA Papers and Proceedings', Vol. 113, pp. 33–38.

Kilpatrick, R. (2022), 'Monitoring sanctions compliance at sea', *Northwestern Journal of International Law and Business* **42**, 221–251.

Lastauskas, P., Proškutė, A. and Žaldokas, A. (2023), 'How do firms adjust when trade stops?', *Journal of Economic Behavior & Organization* **216**, 287–307.

Liberti, J. M., Seru, A. and Vig, V. (2016), ''Information, Credit, and Organization''.

LSEG (2023), DN207488 - Windward Sanction Compliance data to become Fee Liable and Orderable - Customer Action Required, Data Notification.

Norrlöf, C. (2021), Economic statecraft: Finance and Money, Report series, Atlantic Council.

Peress, J. (2014), 'The media and the diffusion of information in financial markets: Evidence from newspaper strikes', *Journal of Finance* **69**(5), 2007–2043.

Ready, R. C. (2018), 'Oil consumption, economic growth, and oil futures: The impact of long-run oil supply uncertainty on asset prices', *Journal of Monetary Economics* **94**, 1–26.

Rosenbaum, P. R. and Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Sun, Y. (2017), 'Do customers learn from stock prices?', *UT Austin Working Paper* .

Van Genugten, J. (2021), 'Conscripting the global banking sector: Assessing the importance and impact of private policing in the enforcement of U.S. Economic sanctions', *Berkeley Business Law Journal* **18**, 136–164.

Wolsing, K., Roepert, L., Bauer, J. and Wehrle, K. (2022), 'Anomaly Detection in Maritime AIS Tracks: A Review of Recent Approaches', *Journal of Marine Science and Engineering* **10**, 1–19.

# Internet Appendix for "The Information Constraint in Market-based Sanctions Enforcement"

Avi Dutt, Abhiroop Mukherjee, George Panayotov, Debjit Roy, Xudong Wen

May 23, 2026

## A Do financial institutions care about sanctions risk: Evidence from an LLM

We focus on the top 50 U.S. financial firms (measured by total assets) from Compustat with a GICS sector code of 40. We download their 10-K reports from SEC websites for 2015-2023. We select paragraphs mentioning "sanction" for textual analysis using GPT-4, where we input the paragraphs and ask:

> **Question**: Does the following paragraph mention economic sanction, for example, enforced by OFAC, and the economic-sanctions-related risk, cost, uncertainty to their business? Output Yes or No, and give a brief reason.

Here are two examples of responses from GPT-4:

> **Input**: for example, the current tensions between the United States and Russia due to the Russian invasion of Ukraine that began in February 2022 and the resulting actions by the United States and a number of other countries in response (including economic sanctions imposed on Russia and the provision of military supplies to Ukraine) could result in retaliatory cyber attacks by Russian threat actors on our business or on third parties with which we do business that have a material impact on our business.
>
> **Response**: Yes, the paragraph mentions economic sanctions imposed on Russia and the resulting potential risks, costs, and uncertainties to the business due to possible retaliatory cyber attacks.

**Input**: governmental regulatory authorities may institute administrative or judicial proceedings that may result in censure, fines, the issuance of cease-and-desist orders, trading prohibitions, the suspension or expulsion of a broker-dealer or member, its officers, registered representatives or employees or other similar sanctions.

**Response**: No, the paragraph does not mention economic sanctions, specifically enforced by OFAC, or the related risk, cost, and uncertainty to their business. It only discusses possible outcomes of administrative or judicial proceedings that can be brought against a broker-dealer or member.

As the examples show, although both inputs mention "sanction," GPT-4 can determine whether it's relevant to our focus. Next, we ask GPT-4 to generate possible reasons for sanctions compliance risk.

**Question**: Please list potential reasons why a firm thinks sanctions compliance is challenging.

Using the 10 reasons provided by GPT-4, we then input the previously identified economic-sanction-related paragraphs for each report and ask the following question:

**Question**: Read the provided firm annual report and mark the reasons from the following list why this firm thinks sanctions compliance is challenging.
1. Complex regulations
2. Frequent updates
3. Multiple jurisdictions
4. Identifying sanctions violators
5. Internal communication
6. Staff training
7. Screening technology
8. Risk assessment
9. Record keeping
10. Enforcement penalties

Give your answer ONLY in a sequence of reason IDs separated by a comma, for example, if your answer is the first 5 reasons, output 1,2,3,4,5

Based on GPT-4's responses, we calculate the percentage of firms related to each reason and report the five most important reasons in Figure 1.

# B Construction of predictors

## B.1 AIS-based predictors

The detailed descriptions for the construction of AIS-based predictors are listed below:

- Identity Change: the indicator equals one if a tanker has changed its MMSI number or flag in the preceding six months, as reported in AIS, and zero otherwise. Windward (2022) has listed identity change as one of "the tried and true techniques of deceptive shipping practices".

- Risky Flag: the indicator equals one if a tanker's flag is among the top 10 that have been associated with moving Iranian and Venezuelan oil (Lloyd's List (2021)). These are the flags of Panama, Cameroon, Vietnam, Djibouti, Cook Islands, Tanzania, Togo, Palau, Russia, and Belize.

- Ship-to-ship Transfer (STS): we construct two types of predictors. The first equals one when the AIS navigation status is recorded as "At Anchor" or "Moored" at locations that are at least $x$ km away from the nearest port, and zero otherwise. We set $x$ to be 10 or 20 (i.e., two versions of this predictor). We obtain port coordinates from the World Port Index (WPI) published by the Maritime Safety Information and supplement this list with estimated ports via a data-driven clustering method, DBSCAN, to reflect ports that are newly built or have been omitted in the WPI. The second equals one when the draft (i.e., the vertical distance between the waterline and the bottom of the hull, which indicates how much cargo the tanker carries) changes by more than one meter while the tanker is at least $x$ km away from the nearest port. We again set $x$ to be 10 or 20 (i.e., two versions of this predictor). We include these predictors as OFAC has listed STS as one of seven deceptive shipping practices to be vigilant against (OFAC (2020)).

- Irregular Trajectory: we define a trajectory as irregular if a tanker's total changes of course within a day are abnormally high. A course is calculated from two adjacent AIS signals, ranging from 0° to 360°, with 0°, 90°, 180°, 270° denoting north, east, south, and west, respectively. A course change is the difference between two adjacent courses, ranging from -180° to 180°. We sum all absolute values of the course changes in a day. To filter out small-scale movements, we calculate the course when the distance between two signals is above a threshold. If the distance doesn't exceed the threshold, we move to the following signals until the requirement is satisfied. We set the distance cutoff to be 1, 2, or 5 km (i.e., three versions). The indicator we use equals one when the

total course changes of a given tanker in a day are above the 95th percentile of the distribution of all such changes within the respective month. We include this predictor following OFAC suggestions (OFAC (2020))

- DBSCAN Outlier: this is a measure aiming to capture anomalous locations conditional on speed and course as identified by DBSCAN (Density-Based Spatial Clustering of Applications with Noise) – a data-driven clustering method that has been employed to detect marine anomalies in a shipping context (e.g., Pallotta et al. (2013)). We first sort the data points into several groups based on speed and course, and then perform DBSCAN within each group, i.e., we cluster conditional on speed and course. The outliers that we find are also conditional, i.e., they deviate in certain ways from tankers to which they are *similar* in other ways. The indicator in this case equals one if a ship is identified as an outlier on a given day, whereby we set the distance parameter of DBSCAN to be 10 or 20 km (i.e., two versions of the predictor).

- Dark Activity: equals one if the time gap between two signals is longer than $T$ days. We set $T = 1, 2, 3$ (i.e., three versions of this predictor). This predictor is motivated by OFAC (2020) and Windward (2022).

## B.2 Satellite-based predictors

Figure A-1 displays three images from the Sentinel-1 satellite as an example, with enlargements showing individual ships at sea. Dfy Graviti – a company specializing in aerospace and maritime AI – helped us process these images to identify and locate ships. Their code used for processing will be made available as part of our replication package. We construct the following satellite-based predictors and provide detailed descriptions for constructing the search area.

- Satellite Detection: We construct the predictors separately for detecting spoofing and dark activity. For spoofing detection, the predictor equals one for day $t$ if a tanker is not seen in its corresponding search area during the period $[t - r, t + r]$. We set $r = 0, 1, 2$ (i.e., three versions) recognizing that spoofing on a given day $t$ may also imply spoofing on adjacent days (e.g., $r$ can be viewed as a smoothing parameter for a time series). For dark activity detection, the predictor equals one if a tanker is not seen in its corresponding search area during its dark period.

We define search areas on satellite images that should contain given tankers, under the null hypothesis that the tankers are not violating sanctions. If a tanker is missing from its search area, we infer that it was likely involved in sanctionable activity. To construct each search

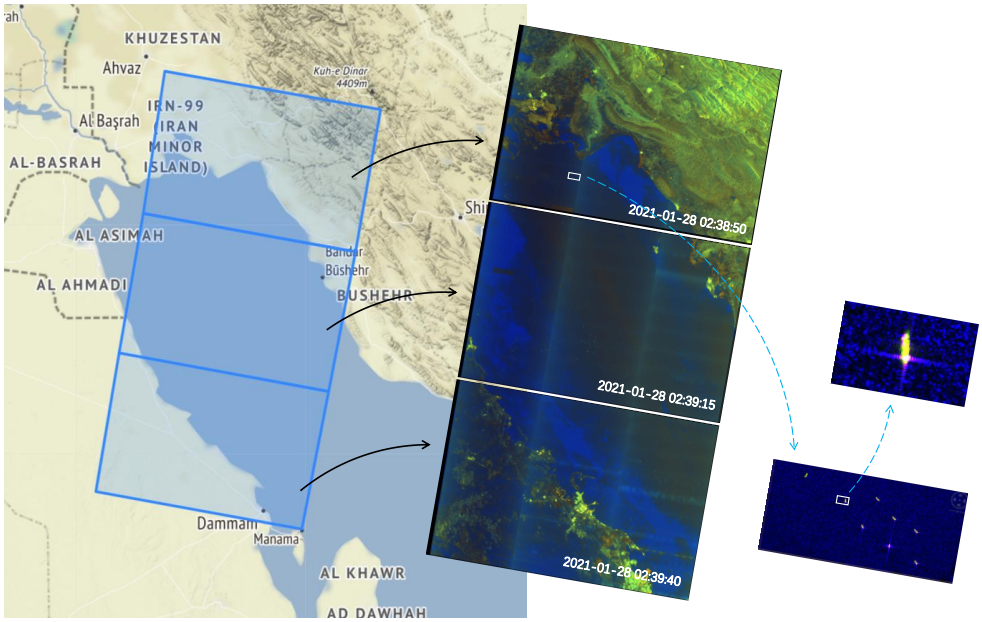



Figure A-1: This figure shows examples of satellite images. These three images were taken a little after 2:30 a.m. on the 28th of January 2021 in the Persian Gulf area. Each image covers a rectangular sector, 250 kilometers wide. The bottom insert on the extreme right shows eight ships as little yellow objects at sea, while the top insert above it shows at the largest enlargement level one of these ships, with individual pixels clearly seen.

area, we estimate its center and size using the two AIS signals closest to the satellite image (just before and after). Using both signals helps narrow the search area.

For spoofing detection, we consider three cases: (I) A stationary tanker, where both adjacent AIS signals are emitted roughly at the same location. Here, the search area is a circle centered around this point. Because a tanker may follow a round-trip route, we only flag a tanker as absent from the area when the time between signals is below a set threshold. (II) A moving tanker with a signal near the satellite image time (within 15 minutes), where the search area centers on this signal. (III) A moving tanker with only distant signals (e.g., a few hours), where we interpolate its trajectory by matching it to others passing through the same area with a similar time span. For dark activity detection, we focus on the first case, using the two adjacent signals just before and after a dark period.

More formally, for a given tanker (call it A) and a given satellite image at time $t^I$, let $t_0$ ($t_1$) denote the time of the closest AIS signal just before (after) $t^I$. Let $x_{t_0}$ and $x_{t_1}$ denote tanker A's locations as per the AIS signals at $t_0$ and $t_1$. For spoofing detection, we construct search areas in the following three cases:

*Case I:* Both $t_0$ and $t_1$ are reasonably close to $t^I$ and the distance between $x_{t_0}$ and $x_{t_1}$ is small. We set the maximum time window to six hours and the maximum distance to five

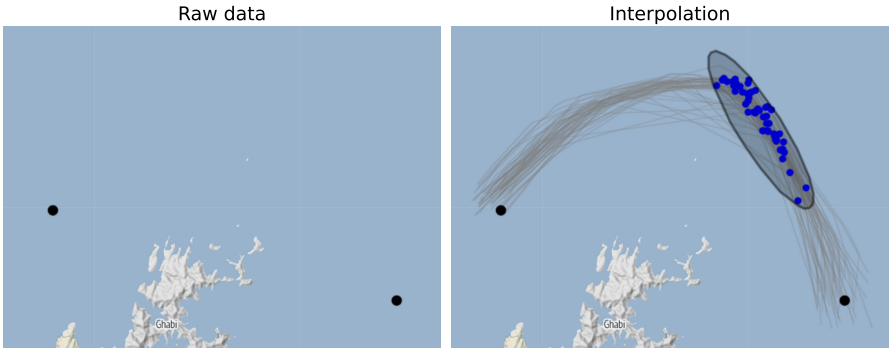



Figure A-2: This figure illustrates interpolation of tanker trajectories in an area seen by a satellite at time $t^I$. The tanker's locations are recorded in AIS at time $t_0 < t^I$ and $t_1 > t^I$, with no signals in-between (we consider only cases when both of these times are within six hours of $t^I$ to limit noise)). The black dots $(x_0)$ and $(x_1)$ denote the tanker's locations at $t_0$ and $t_1$. We determine the area to be searched in the satellite image by interpolating trajectories (shown by grey lines) of comparable tankers that passed near these points in similar time frames and directions. Then we find the locations (shown as blue dots) of these comparable tankers at $t^I$. We define the search area as the smallest ellipse containing all blue dots with 95% probability.

kilometers. The search area is defined as a circle centered at the midpoint between $x_{t_0}$ and $x_{t_1}$ with a diameter equal to $\max(|x_{t_0} - x_{t_1}|, 0.5)$ km.

*Case II:* One of $t_0$ or $t_1$ is over six hours from $t^I$, but the other is within 15 minutes. We define the search area as a circle centered on the closer signal's location. The radius assumes the ship could travel at a maximum speed of 15 knots (28 km/h) between the signal and $t^I$, with a minimum radius as in Case I.

*Case III:* Both $t_0$ and $t_1$ are again reasonably close to $t^I$ (within six hours), but the distance between $x_{t_0}$ and $x_{t_1}$ is large. In this case, we define the search area based on a data-driven interpolation method. The main idea is to interpolate the location of tanker A using the trajectories of other tankers that have passed near both $x_{t_0}$ and $x_{t_1}$, in the same direction and over a similar time interval.

Figure A-2 illustrates Case III, where we match tanker A to comparable tankers by location and average speed. We collect the location of each comparable tanker at its time corresponding to $t^I$ for tanker A (the blue dots in the figure). We define the search area as the smallest ellipse that contains all of these collected locations, except those that are more than five standard deviations away from the center of the interpolation locations. We use ellipses to reduce the size of the search area.

To be included in the matching, we require a tanker to have sent a signal within a five-kilometer distance from $x_{t_0}$, and then another one within a five-kilometer distance from $x_{t_1}$, whereby the time it traveled between these two signals is within 30 minutes of the difference

$t_1 - t_0$. To increase the number of matches, we search for comparable tankers in the three-month period between December 2020 to February 2021. We also validate our method by sequentially leaving out signals as an out-of-sample test and show that our interpolation method yields a successful coverage for 96% of the test cases.

When detecting dark activity, Cases II and III are not applicable, since by design the time interval between the adjacent AIS signals is larger than a few hours – we have defined a dark period to be at least 24 hours long. Therefore, we only consider here Case I, but modify it so that $t_0$ is the last signal before a dark period, and $t_1$ is the first signal after a tanker emerges from the dark period.

Finally, note that alternative detection methods, such as identifying ships at port, are infeasible because satellite images of standard resolutions cannot distinguish between multiple vessels of similar size and shape. Furthermore, sanctions violators sometimes have altered their vessels' appearance (IHS-ACSS, 2022).

### B.3 Summary Statistics

Table A-1 presents summary statistics for variables used in the detection of Spoofing and Dark Activity.

## C Detecting sanctions violators

### C.1 Predictive Models and their Performance

To detect spoofing, we create a tanker-day sample where the dependent variable is one if a tanker is observed in Iran (per our ground truth dataset), but its AIS signals indicate another location; it is zero otherwise. To detect dark activity, our dependent variable is one if a tanker is observed in Iran during a dark period exceeding 24 hours, and zero otherwise.

To optimally combine predictors, we employ decision trees and neural networks with 10-fold cross-validation to fine-tune model parameters; the sample is randomly divided into 10 sub-samples, models are trained with nine sub-samples, leaving one for validation. This procedure is repeated 10 times to select the best hyper-parameter combination based on McFadden's pseudo-$R^2$. Detection performance is evaluated from the best cross-validated model using its pseudo-$R^2$.

Table A-2 presents our results. Panel A reveals that our best AIS-based detection models achieve a pseudo-$R^2$ of about 10% for detecting spoofing and dark activity. Incorporating satellite data enhances detection only modestly, from 10% to 16%. Panel B evaluates the incremental contribution of each predictor category. While overall pseudo-$R^2$ values remain

Table A-1: Summary statistics of detection samples

This table reports the summary statistics for the detection samples in Section 3. We separately detect spoofing and dark activity. Panel A reports the sample sizes and Panel B reports the mean and standard deviation of predictors. The construction of the predictors is described in Section 3.3. "-" in Panel B indicates that the sample does not include the respective predictor.

Panel A: Sample Size

| | Spoofing Detection | Dark Detection |
|---|---|---|
| No. observations | 17,496 (tanker-day) | 1,552 (dark periods) |
| No. unique tankers | 1,481 | 746 |
| No. unique violators | 17 | 20 |

Panel B: Predictors

| | Spoofing Detection | | Dark Detection | |
|---|---|---|---|---|
| | Mean | Std | Mean | Std |
| Identity Change | 0.118 | 0.322 | 0.144 | 0.352 |
| Risky Flag | 0.263 | 0.440 | 0.288 | 0.453 |
| Ship-to-ship Transfer | | | | |
| • anchor 10 km away from ports | 0.218 | 0.413 | 0.363 | 0.481 |
| • anchor 20 km away from ports | 0.081 | 0.273 | 0.137 | 0.344 |
| • draft change 10 km away from ports | 0.036 | 0.186 | 0.151 | 0.358 |
| • draft change 20 km away from ports | 0.027 | 0.161 | 0.119 | 0.324 |
| Irregular Trajectory | | | | |
| • distance cutoff = 1 km | 0.050 | 0.218 | 0.053 | 0.225 |
| • distance cutoff = 2 km | 0.050 | 0.218 | 0.068 | 0.251 |
| • distance cutoff = 5 km | 0.050 | 0.218 | 0.079 | 0.269 |
| DBSCAN Outlier | | | | |
| • distance parameter = 10 km | 0.131 | 0.338 | 0.318 | 0.466 |
| • distance parameter = 20 km | 0.049 | 0.216 | 0.142 | 0.349 |
| Dark Activity | | | | |
| • at least 2-day dark period | - | - | 0.406 | 0.491 |
| • at least 3-day dark period | - | - | 0.244 | 0.429 |
| Satellite Images | | | | |
| • 1-day window | 0.010 | 0.101 | - | - |
| • 3-day window | 0.028 | 0.165 | - | - |
| • 5-day window | 0.042 | 0.200 | - | - |
| • during dark period | - | - | 0.025 | 0.157 |

low, the Dark Activity and Ship-to-ship Transfer predictors are most important here, with incremental pseudo-$R^2$ of 9.4% and 6.2%, respectively.

Overall, we find that the detection of sanctions violators is challenging, even when combining numerous predictors mentioned in industry and regulatory sources, and using sophisticated methods/models. It is possible that governments and certain corporations with superior data access (e.g., Starlink) do not face these issues, but most third parties lack access to such data; and yet they are still tasked with such detection to comply with sanctions

Table A-2: Performance of ML detection models

In Panel A, this table shows detection accuracies for spoofing and dark activity, and their simple average ("Total"), using either AIS data alone, or together with satellite data. In Panel B, $\Delta$pseudo-$R^2$ for a given category is the difference between the pseudo-$R^2$ of a full model and the same model that sets this particular category to zero (note that all our predictors are zero-one indicators). Negative pseudo-$R^2$ or $\Delta$pseudo-$R^2$ are treated here as zero. We report simple averages of the respective $\Delta$pseudo-$R^2$'s across the tree and neural network models, and spoofing and dark detection. In Panel C, we report the total number of tankers that violated Iranian sanctions during January 2021 (as per our proprietary dataset), the number of tankers among them labeled as high/moderate risk in Windward's list, and the number of tankers detected by our ML models. Results are shown for different confidence levels ($cl = 99\%$, $95\%$, and $90\%$).

Panel A: Model Performance

| | | Model performance (pseudo-$R^2$) | |
|---|---|---|---|
| Detect | Model | Using only AIS data | Using AIS and satellite data |
| Total | Tree | **9.1%** | **15.2%** |
| | NN | **10.2%** | **16.3%** |
| Spoofing | Tree | 6.7% | 11.8% |
| | NN | 7.2% | 13.8% |
| Dark | Tree | 11.5% | 18.7% |
| | NN | 13.3% | 18.8% |

Panel B: Predictor Importance

| Category | $\Delta$pseudo-$R^2$ |
|---|---|
| Satellite Detection | 6.2% |
| Identity Change | 3.7% |
| Risky Flag | 2.8% |
| Irregular Trajectory | 0.2% |
| Ship-to-ship Transfer | 6.2% |
| DBSCAN Outlier | 5.3% |
| Dark Activity | 9.4% |

Panel C: Comparison between the Windward's list and ML-model detection

| | |
|---|---|
| Total Violators: | 33 |
| on Windward's list: | **27** |

| | | | Number of violators detected | | |
|---|---|---|---|---|---|
| Detect | Model | Predictors | $cl = 99\%$ | $cl = 95\%$ | $cl = 90\%$ |
| Total | Tree | Satellite + AIS | **5** | **12** | **17** |
| | NN | Satellite + AIS | **5** | **13** | **15** |
| | Tree | Only AIS | 3 | 8 | 14 |
| | NN | Only AIS | 0 | 5 | 13 |

rules. It is also possible that our results reflect limitations in our ground truth dataset or models, rather than a general difficulty in detection for third party market participants – in Section C.2, we provide a discussion and further evidence on such detection being generally challenging, even beyond our data or models.

## C.2 Is detection truly difficult?

One might question whether our results indicate a general difficulty in detection or merely reflect limitations in our ground truth dataset or models.

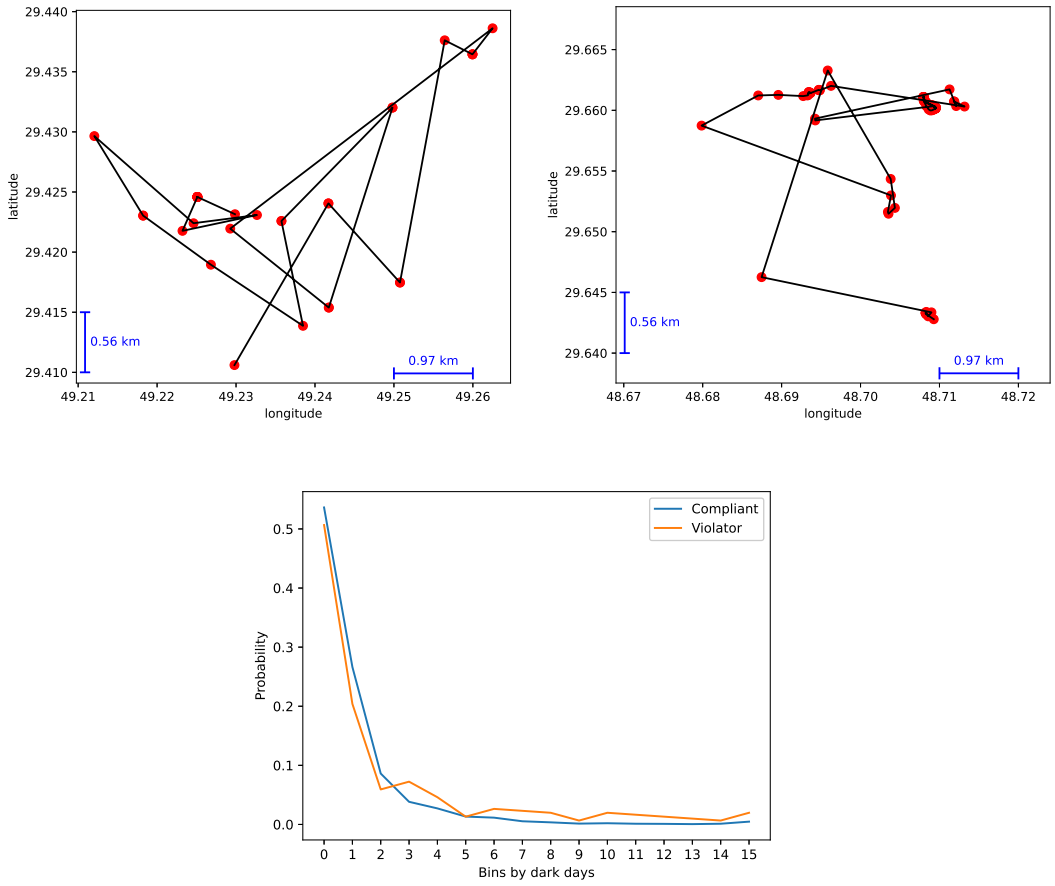



Figure A-3: The top-left (top-right) plot shows irregular trajectories for a sanctions-violating tanker from our proprietary dataset (a sanctions-compliant tanker). The axes of these two plots show degrees of latitude and longitude, and the red dots show the coordinates of each AIS signal emitted by each tanker. The bottom plot shows the distribution of dark days in Jan 2021 for tankers entering the Persian Gulf, conditional on the tanker going dark for at least half a day. The $x$-axis shows bins of dark time: bin 0 is from 12 hours to 24 hours; bin $x$ ($x = 1, 2, \cdots, 14$) is from $x$ to $x + 1$ days; bin 15 is equal or greater than 15 days. The blue (orange) line represents sanctions-compliant (sanctions-violating) tankers.

First, a closer examination of our data reveals that sanctions-evading predictors exhibit similar patterns for both violating and compliant tankers. For example, as shown in Figure A-3, compliant tankers (top-right) can show irregular trajectories similar to those of violators (top-left), and the dark time distributions for both tanker types overlap significantly. Legitimate factors like weather, mechanical issues, or AIS data errors (Weng et al. (2022))

contribute to this low signal-to-noise ratio, compounded by sparse satellite data.

Second, evidence of detection challenges abounds, even beyond our data and methods. For instance, only 10 of the 33 tankers from our Iran-sanctions violators dataset appear on UANI's "The Ghost Armada" list, as of August 2023, illustrating the broader difficulty in identifying violators.

Overall, detecting sanctions-violating tankers is challenging for third party market participants. While governments and certain corporations with superior data access (e.g., Starlink) may not face these issues, most third parties lack access to such data but are still required to comply.

# D The Refinitiv Disclosure

## D.1 Windward's sanction risk data

Figure A-4 shows a screenshot from Refinitiv Eikon illustrating Windward's sanctions-risk measures for one tanker (A STAR). We report the sample filtering process and the sample size, i.e., the number of tankers, after each step, in Table A-3. Summary statistics on tankers included in the Windward dataset are presented in Table A-4. This table also presents balance tests on the matched sample of tankers.

Table A-3: Sample filtering

This table reports the sample filtering process and the sample size after each step, i.e., the number of tankers.

| Data cleaning steps | # tankers |
|---|---|
| Full tanker list from Reuters | 13,997 |
| Keep oil tankers | 10,322 |
| Keep large tankers (i.e., Medium, Panamax, Aframax, Suezmax, VLCC) | 5,466 |
| Drop tankers that were sanctioned before the disclosure | 5,392 |
| Drop tankers without AIS data (i.e., unable to calculate propensity score) | 5,363 |
| Reg. sample 1 (merge to fixture data) | 3,628 |
| Reg. sample 2 (merge to fixture data and require fixture type is WS) | 2,852 |
| Reg. sample 3 (merge to ownership data) | 3,312 |

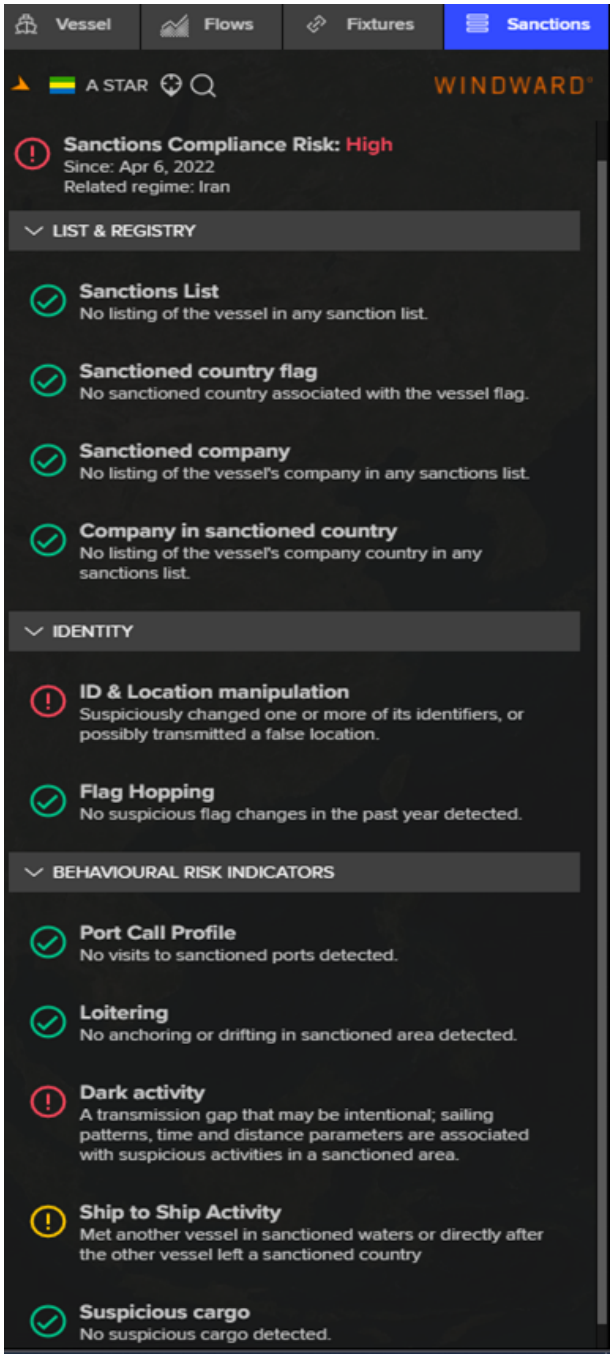



Figure A-4: This figure shows a screenshot from Refinitiv Eikon illustrating Windward's sanctions-risk measures for one tanker (A STAR). The tanker is classified as high-risk with respect to the Iranian sanctions, because Windward has found strong evidence that it has changed its ID and/or has been involved in spoofing, and also has had long dark periods (transmission gaps) that can be associated with suspicious activities. Windward has also found some (weaker) evidence that the tanker was involved in ship-to-ship transfers.

Table A-4: Summary statistics of Windward's tankers sample

This table reports the summary statistics for the sample of tankers we use to examine the effects of the disclosure event. Panel A reports sample sizes and tanker classification. Panel B reports the balance tests on tanker characteristics between treated (i.e., high- and moderate-risk) and control (i.e., low-risk) groups. To construct the control group, we conduct propensity score matching (PSM) for each treated tanker within its tanker type, i.e., we require matched tanker(s) to have the same tanker type. The propensity score is calculated by regressing the high/moderate-risk tanker indicator on pre-period tanker characteristics using ML methods: average outputs from a decision tree and a neural network. The variables used to construct the propensity score are the same as the list in Panel B. We calculate matching weights based on a Gaussian kernel with a bandwidth of 0.05. The variable differences and corresponding $t$-value are reported. For this tanker-level regression, standard errors are clustered at the flag level. Panel C reports sample averages of freight rates (or fixtures).

Panel A: Tanker sample

| No. unique tankers | 5,363 | | | |
|---|---|---|---|---|
| | Sanction risk | | | |
| Tanker type | Low | Moderate | High | Total |
| Medium | 1,579 | 423 | 306 | 2,308 |
| Panamax | 297 | 88 | 55 | 440 |
| Aframax | 533 | 295 | 280 | 1,108 |
| Suezmax | 309 | 179 | 167 | 655 |
| VLCC | 672 | 64 | 116 | 852 |
| Total | 3,390 | 1,049 | 924 | 5,363 |

Panel B: Tanker characteristics balance test

| | Treated (High/Mod.-risk) | Control (Low-risk) | Diff. | $t$-value |
|---|---|---|---|---|
| Identity Change | 0.169 | 0.165 | 0.003 | [0.12] |
| Risky Flag | 0.276 | 0.264 | 0.012 | [0.24] |
| Ship-to-ship Transfer | | | | |
| • anchor 10 km away from ports | 2.652 | 2.733 | -0.082 | [-1.61] |
| • anchor 20 km away from ports | 1.424 | 1.475 | -0.051 | [-0.67] |
| • draft change 10 km away from ports | 1.620 | 1.668 | -0.048 | [-1.63] |
| • draft change 20 km away from ports | 1.297 | 1.334 | -0.037 | [-0.98] |
| Irregular Trajectory | | | | |
| • distance cutoff = 1 km | 1.087 | 1.064 | 0.024 | [0.40] |
| • distance cutoff = 2 km | 1.426 | 1.414 | 0.012 | [0.23] |
| • distance cutoff = 5 km | 1.420 | 1.434 | -0.014 | [-0.27] |
| DBSCAN Outlier | | | | |
| • distance parameter = 10 km | 3.744 | 3.670 | 0.075 | [1.33] |
| • distance parameter = 20 km | 3.023 | 2.946 | 0.077 | [1.31] |
| Dark Activity | | | | |
| • at least 2-day dark period | 0.883 | 0.872 | 0.010 | [0.27] |
| • at least 3-day dark period | 0.600 | 0.589 | 0.011 | [0.31] |

Panel C: Average fixture

| | Average fixture in July 2023 (one month before the event) | | |
|---|---|---|---|
| Full sample | Low risk | Moderate risk | High risk |
| 165.3 | 129.8 | 135.2 | 126.8 |

## D.2 Attention to Windward

Figure A-5 shows page views of the Windward.AI homepage. It increased dramatically right after the list was put up on Reuters, suggesting that market participants likely paid attention to this disclosure.

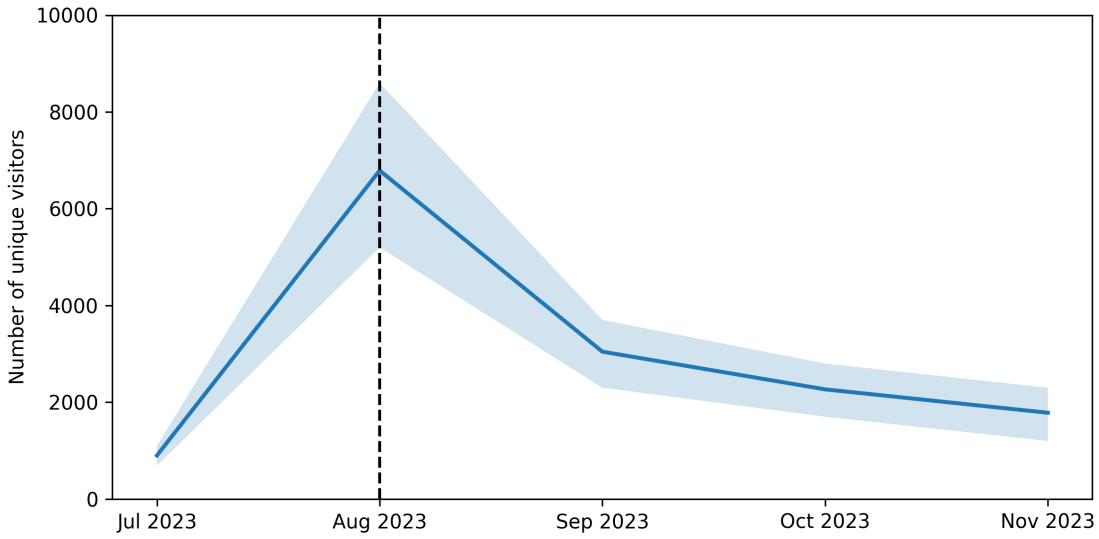



Figure A-5: This figure plots the number of unique visitors of the domain https://windward.ai/ (and its sub-domains), from July 2023 to November 2023. The data is extracted from Semrush, a competitive analysis platform operated by Semrush Holdings, Inc. The shadowed area represents the estimation deviation (i.e., the error band provided by the platform). The vertical dashed line indicates the event month of the Refinitiv disclosure of high/moderate-risk tankers.

## D.3 Optimal mix of public-data-based detection models and Windward's List

Figure A-6 compares the out-of-sample predictability of future sanctions achieved using weighted averages of Windward's and public-data-based predicted outcomes with various weights.

## D.4 Predicting suspected violators in the list of Bai et al. (2025)

We examine the predictive accuracy of our ML models for the list of 25 suspected violators used in Bai et al. (2025). Following the same approach as when using our ground truth dataset, we employ decision trees and neural networks with 10-fold cross-validation to train models and then examine out-of-sample predictability. Inferring from the 25 tankers shown

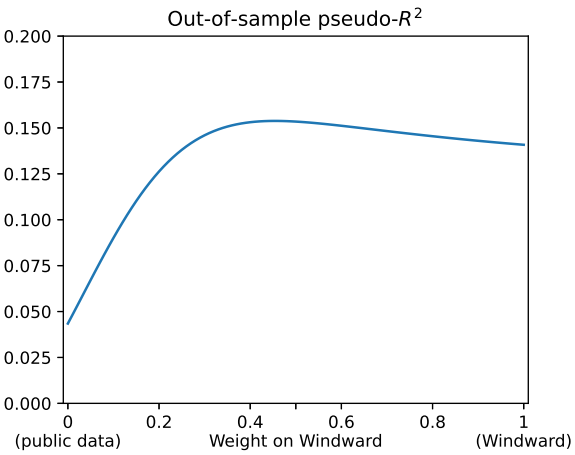



Figure A-6: This figure compares the out-of-sample predictability of future sanctions achieved using weighted averages of Windward's and public-data-based predicted outcomes with various weights. We plot the corresponding out-of-sample performance measures against the weight for Windward's list. Windward's predicted outcome is a variable equal to one for high-risk tankers and zero otherwise. For the public-data-based predicted outcome, we take the view of a third party predicting future sanctioned tankers at the end of July 2023 (i.e., right before the Refinitiv disclosure). For our training sample, the predictors are constructed from data from August 2022 till January 2023 and the dependent variable is a zero-one indicator of tankers sanctioned between August 2022 and July 2023. The public data include AIS-based predictors and satellite-based predictors. We take the average of the ML models (i.e., decision trees and neural networks). The public-data-based predicted outcome is the estimated probability of a tanker being sanctioned after July 2023. The out-of-sample evaluation does not include tankers that were already sanctioned at the prediction time, i.e., the end of July 2023.

in Table 1 in Bai et al. (2025), we focus on the three largest tanker types (VLCC, Suezmax and Aframax), which account to a total of about 2600 tankers in our sample (comparable to their total of 2160). Among these, we select the 558 with the highest probability of being violator averaged across our ML models – this number corresponds to the "dark" tankers that they identify. Our 558 most-likely violators include 19 of their 25 suspects. Therefore, similar to their evaluation measure, our sample yields a precision of 76%, which is very close to the 83% that they report.

# E Estimating the effects of the Refinitiv disclosure

The key issue is that the untreated tankers, i.e., those classified as low-risk by Windward, may not provide accurate counterfactuals for high/moderate-risk tankers. This is because the low-risk tankers could be so different from the high/moderate-risk tankers that their fixtures would have trended differently from those of the high/moderate-risk tankers even without

the disclosure. In that case, the conventional difference-in-differences (DiD) estimator would be biased, because its key identification assumption – parallel trends for the treatment and control – no longer holds.

We use two methods to address this issue. The first is propensity score matching (PSM), introduced by Rosenbaum and Rubin (1983) to construct counterfactuals relying on the strong ignorability assumption ("unconfoundedness"). PSM is an intuitive method and makes it easy to visualize the data using standard DiD plots. Second, we follow Abadie (2005) and employ a semiparametric DiD estimator. Abadie's estimator only requires conditional mean independence – that is, conditional on covariates, the expected value of the outcome variable does not depend on treatment status.[1] Moreover, PSM requires *estimating* matching weights as a function of propensity scores, choosing the functional form. In contrast, Abadie (2005)'s method offers lower degrees of researcher freedom regarding specification choice.

Consider the following setup to fix ideas in our context. Let $t = 0$ ($t = 1$) denote the pre-treatment period (post-treatment period), and $X$ is a vector of tanker observed characteristics. Let $D = 1$ ($D = 0$) indicate that the tanker is treated (untreated), i.e., it is classified as high/moderate risk (low-risk) in the Windward disclosure. Let $Y_t$ be the tanker's observed fixture at period $t$. Further, suppose $Y_1^T|D = 1$ is the fixture for a treated tanker at $t = 1$, and $Y_1^T|D = 0$ is its counterfactual fixture – i.e., what its fixture would have been at $t = 1$ had it not been treated. Similarly, $Y_1^U|D = 0$ denote the fixture for an untreated tanker at $t = 1$, and $Y_1^U|D = 1$ denote its counterfactual fixture – i.e., what its fixture would have been at $t = 1$ had it actually been treated. The quantity of interest is the average treatment effect on the treated (ATT):

$$ATT = E\left[Y_1^T - Y_1^U|D = 1\right], \tag{A-1}$$

which is the difference between the fixture for a disclosed-as-high/moderate-risk tanker and what its fixture would have been, had it not been disclosed as high/moderate risk. When there is a need to indicate a particular tanker, e.g., tanker $i$, we add a subscript $i$ to the variables, e.g., $X_i$ and $Y_{i,t}$. In the cases without confusion, we drop the subscript of $i$ for simplicity.

## E.1 Propensity score matching (PSM)

PSM addresses the challenge that the conventional DiD identification assumption

$$E\left[Y_1^U - Y_0|D = 1\right] = E\left[Y_1^U - Y_0|D = 0\right], \tag{A-2}$$

[1] See, e.g., Heckman et al. (1997) and Abadie (2005).

may not hold when the differences between the characteristics of treated and control tankers, i.e., between $X|D=1$ and $X|D=0$, affect the dynamics of outcome variables. For every treated tanker $i$, PSM tries to find tanker(s) $j$ which have very *similar* characteristics $X_j$ to those of tanker $i$ but were actually classified by Windward as low-risk; then it treats such tanker(s) $j$ as counterfactual. Rosenbaum and Rubin (1983) show that, under unconfoundedness (i.e., $Y_1^T, Y_1^U \perp\!\!\!\perp D \mid X$), it is sufficient to focus solely on tankers with similar propensity $p = \Pr(D=1|X)$ to have been labeled as high/moderate risk ex-ante. That is, we can define 'similar' tankers relying only on dimensions of $X$ that matter for risk labeling, rather than considering matching on the full vector of characteristics $X$.

In our notation, PSM replaces $[Y_{i,1}^U - Y_{i,0}|D_i=1]$, i.e., the counterfactual fixture changes of high/moderate-risk tanker $i$ in the absence of treatment, with $\sum_{j:D_j=0} w(j,i)(Y_{j,1}-Y_{j,0})$, i.e., the weighted average of observed fixture changes of low-risk tanker(s) $j$ that had a similar ex-ante propensity to be labeled as high/moderate risk. $w(j,i)$ is determined by comparing $p_i$ and $p_j$ (Rosenbaum and Rubin, 1983). specifically, we give a higher weight for a low-risk tanker $j$ if $p_j$, the conditional probability of being labeled as high/moderate risk based on $X_j$, is closer to that of the target tanker $i$. Once the weights $w(j,i)$ of matching tankers are obtained, ATT can be estimated as follows:

$$\hat{ATT} = \frac{1}{N_1} \sum_{i:D_i=1} \left[ (Y_{i,1} - Y_{i,0}) - \sum_{j:D_j=0} \hat{w}(j,i)(Y_{j,1} - Y_{j,0}) \right], \tag{A-3}$$

where $N_1$ is the total number of treated tankers and $\hat{w}(j,i)$ is the estimated value of $w(j,i)$. ATT averages the differences between the fixture changes around the disclosure event for treated tankers and for the matched untreated tankers.

## E.2 Abadie's (2005) semiparametric difference-in-differences

Abadie (2005) relies on the following identification assumption:

$$E\left[Y_1^U - Y_0 | X=x, D=1\right] = E\left[Y_1^U - Y_0 | X=x, D=0\right]. \tag{A-4}$$

This conditional identification assumption is more suitable to our setting because it is more plausible that conditional on the same tanker size, AIS emission pattern, trajectories, ownership structure, etc., the fixture of a disclosed high/moderate-risk tanker in the absence of disclosure would follow parallel dynamics with that of a low-risk tanker having the same value of all these characteristics in $X$. To proceed, Abadie (2005) shows that ATT defined in equation (A-1) can be written as an integral along all dimensions of $X$: $E[...|D=1] = \int E[...|X=x, D=1]\, dP(X=x|D=1)$. After substituting into the identi-

fication assumption in equation (A-4) and some algebraic manipulation, ATT equals to

$$ATT = \frac{1}{\Pr(D=1)} \cdot E\left[\frac{D-p}{1-p} \cdot (Y_1 - Y_0)\right], \tag{A-5}$$

where $p = \Pr(D=1|X)$, as defined before, is the propensity score. To gain some intuition, we can partition the expectation in equation (A-5) based on $D=0$ and $D=1$, which gives

$$ATT = E\left[Y_1 - Y_0 | D=1\right] - E\left[\frac{p/\Pr(D=1)}{(1-p)/(1-\Pr(D=1))} \cdot (Y_1 - Y_0)|D=0\right]. \tag{A-6}$$

In equation (A-6), the first term is the average of changes of observed fixtures for risky tankers, and the second term is a propensity-score-based-weighted average of changes of observed fixtures for low-risk tankers. Under the perfect randomized treatment assignment, $p = \Pr(D=1|X) = \Pr(D=1)$ and the propensity-score-based weight is always one, which returns back to the conventional DiD estimator. When $p$ differs from $\Pr(D=1)$, the method will give higher (lower) weights to the low-risk tankers with higher (lower) $p$, i.e., those who are closer (not close) to the high/moderate-risk tankers in terms of propensity scores.

## E.3 How are fixtures measured?

In shipping contracts, spot fixture rates are often quoted relative to "Worldscale" (WS), i.e., a baseline rate calculated for a standardized vessel on a round-trip voyage between two specified ports (or between ports in two specified areas). Such baseline rates are defined as "WS100" and an actual quote can be given relative to the baseline – for example, "WS75" or "WS200" would denote 75% of the baseline rate or double this rate for a charter between the baseline ports. WS rates are revised annually based on updated bunker prices, port costs, and exchange rates.[2] The key advantage of the Worldscale is that it makes freight rates across different contracts more comparable. About two thirds of the contracts in our sample are in WS format, and we focus our analysis on WS rates.

## E.4 Robustness checks

We consider the following robustness checks on our fixture changes estimation. First, we use propensity scores from a logistic regression, rather than the baseline machine learning models. Second, we drop the fixtures in August 2023 to allow about two weeks for the disclosure effect to come into force. Third, we drop the fixtures with Russian ports, i.e., ports of Novorossiysk, St. Petersburg, Primorsk, Tuapse, and Taman in our data, to check that our effects are not coming solely from Russia-related changes in maritime activity around

[2] See https://www.worldscale.co.uk/ for further details.

Table A-5: Robustness checks

This table reports on robustness checks for the fixture changes estimation. The sample and construction of propensity score are the same as in Table 3, except for the features indicated in the first row of the table. In Panel A, PSM uses a Gaussian kernel with a bandwidth of 0.01. Standard errors are double-clustered at the tanker and time×tanker-type levels (except for the last column). In Panel B, we manually subtract the cross-sectional mean by tanker type to account for time×tanker-type fixed effects, since Abadie's estimator is derived without directly accounting for macro trends through time fixed effects. Standard errors are as derived in Abadie (2005) (except for the last column). *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

| | Propensity score from logistic reg. | Drop Aug 2023 | Drop Russian ports | Bootstrap s.e. (500 replications) |
|---|---|---|---|---|
| Panel A: PSM-DiD | | | | |
| | Dependent Variable: Fixtures | | | |
| ATT for High Risk | -10.516* | -15.675** | -10.999** | -16.063*** |
| | [-1.70] | [-2.51] | [-2.08] | [-3.54] |
| Obs. (tanker-month) | 5,383 | 4,957 | 5,200 | 5,378 |
| Time × Tanker Type FE | Yes | Yes | Yes | Yes |
| Tanker FE | Yes | Yes | Yes | Yes |
| Panel B: Abadie (2005) Semiparametric DiD | | | | |
| | Dependent Variable: Fixtures demeaned by time × tanker type | | | |
| ATT for High Risk | -12.280*** | -14.854*** | -9.945** | -13.749*** |
| | [-2.87] | [-3.29] | [-2.26] | [-3.20] |
| Obs. (tanker) | 1,089 | 976 | 1,037 | 1,036 |

this time. Finally, we use bootstrap standard errors with 500 replications instead of clustered standard errors. The results are shown in Table A-5.

## E.5 Changes in charterers' behavior

Did U.S. and U.S.-allied charterers start avoiding tankers classified as high or moderate risk by Windward? We obtain charterers' names from fixtures data and determine the countries of their headquarters via LinkedIn or company websites. U.S.-allied charterers include companies from the UK, European Union, Australia, New Zealand, or Japan. The dependent variable is the number of tankers of a given risk type that each charterer employs each month. The regression specification is

$$\#tankers_{h,t} = \left[\beta^U \times US_charterer_h + \beta^A \times US_allied_charterer_h\right] \times \mathbb{I}_{\{t\geq 0\}} + \alpha_h + \gamma_t + \epsilon_{h,t}, \quad \text{(A-7)}$$

where $h$ denotes a charterer.

Table A-6 presents the estimation results for Eq.(A-7). Similar to our previous results,

Table A-6: Difference-in-differences analysis of changes of charterers

In this table, we examine whether U.S. or U.S.-allied companies tend to stop chartering high/moderate-risk tankers after the Refinitiv disclosure. U.S.-allied companies are those from the UK, European Union, Australia, New Zealand, or Japan. The dependent variable is the number of tankers of a given risk type that each charterer employs in a month. The sample period is from Feb 2023 to Mar 2024. The post-period starts from Aug 2023. We regress the dependent variable on the indicators of U.S. and U.S.-allied charterers interacted with the post-period indicator controlling for charterer fixed effects and time fixed effects. The coefficients of the DiD estimation are reported. The standard errors are double clustered at the charterer and time level. *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

| | Dependent Variable: No. tankers for each charterer-month | | |
|---|---|---|---|
| | No. high-risk tankers | No. mod.-risk tankers | No. low-risk tankers |
| U.S. charterer $\times \mathbb{I}_{\{t\geq 0\}}$ | -0.358* | -0.131 | 0.065 |
| | [-1.91] | [-0.67] | [0.09] |
| U.S.-allied charterer $\times \mathbb{I}_{\{t\geq 0\}}$ | 0.013 | -0.030 | 0.136 |
| | [0.08] | [-0.12] | [0.31] |
| Obs. (charterer-month) | 1,820 | 1,820 | 1,820 |
| Time FE | Yes | Yes | Yes |
| Charterer FE | Yes | Yes | Yes |

we find a significant drop in high-risk tankers usage by U.S. charterers, while effects for U.S.-allied charterers are not statistically significant. For perspective, the average U.S. charterer employed 0.652 high-risk tankers in July 2023, and the estimated 0.358 drop represents a 54.9% reduction (0.358/0.652).

## E.6 Selection into dataset

One question related to our estimated treatment effects could be whether there are potential selection issues: the information shock may affect the high/moderate-risk tankers' incentive to report their fixture contracts to the data vendor, from which we source our numbers. For example, if after the information becomes public, some high/moderate-risk tankers are reluctant to disclose their contracts, there could be a concern that our estimation is biased by the selection issues. We test this hypothesis by using as our dependent variable the count of the number of fixture contracts per month for each tanker, using the same specification as our fixture results. To enter into the sample, we require a tanker to report at least one fixture record during our sample period.

If there are selection issues, we will observe that the number of fixture contracts for high/moderate-risk tankers significantly decreases if they want to avoid disclosing, after Windward publishes their list. The results in Table A-7 show that all coefficients are both statistically and economically insignificant for either of our two methods.

This suggests that the owners of high/moderate-risk tankers do not seem to change their reporting of fixture contracts to the dataset, so such selection issues do not seem to be important.

Table A-7: Difference-in-differences analysis of sample selection

This table examines whether there are potential selection issues by estimating the average treatment effects on the treated (ATT) on the number of fixture records for high/moderate-risk tankers after the Refinitiv disclosure using PSM-DiD and Abadie (2005)'s semiparametric DiD. The dependent variable is the number of fixture records for each tanker-month, and it is set to zero if there are no fixture records. We exclude the tankers with all zero values of the dependent variable. The sample period is from Feb 2023 to Mar 2024. The post-period starts from Aug 2023. The propensity score is calculated by regressing the high/moderate-risk tanker indicator on pre-period tanker characteristics using ML methods. The variables used to construct the propensity score are shown in Table A-1. In Panel A, we match tankers within each tanker type and calculate weights based on propensity scores and a Gaussian kernel with a bandwidth of 0.01, 0.03, or 0.05. Then, we do DiD estimation in the matched sample, controlling for tanker fixed effects and time×tanker-type fixed effects. The standard errors in Panel A are double clustered at the tanker and time×tanker-type levels. In Panel B, we implement Abadie's method. The standard errors in Panel B are as derived in Abadie (2005). *, **, *** denote significance at the 10%, 5%, and 1% level, respectively.

| Panel A: PSM-DiD | | | |
|---|---|---|---|
| | Dependent variable: No. fixture records in each month | | |
| | Bandwidth = 0.01 | Bandwidth = 0.03 | Bandwidth = 0.05 |
| ATT for High Risk | 0.020 | 0.013 | 0.011 |
| | [0.68] | [0.48] | [0.43] |
| Obs. (tanker-month) | 37,338 | 37,338 | 37,338 |
| ATT for Moderate Risk | 0.009 | 0.002 | 0.002 |
| | [0.45] | [0.11] | [0.10] |
| Obs. (tanker-month) | 41,776 | 41,776 | 41,776 |
| Time × Tanker Type FE | Yes | Yes | Yes |
| Tanker FE | Yes | Yes | Yes |
| Panel B: Abadie (2005) Semiparametric DiD | | | |
| | Raw no. records | No. records demeaned by time | No. records demeaned by time × tanker type |
| ATT for High Risk | 0.009 | 0.009 | 0.015 |
| | [0.45] | [0.47] | [0.78] |
| Obs. (tanker) | 2,447 | 2,447 | 2,447 |
| ATT for Moderate Risk | -0.003 | -0.003 | 0.000 |
| | [-0.18] | [-0.19] | [0.02] |
| Obs. (tanker) | 2,984 | 2,984 | 2,984 |

# F Details related to the model

## F.1 Signals, risk labels, and rational expectations

A key component of our model is that tankers and Clean exporters form rational expectations of the authority's signal from risk labels. We calculate the learning process (i.e., $\mathbb{E}[a|G,L]$, $\mathbb{E}[a|G,H]$, $\mathbb{E}[a|B,L]$, and $\mathbb{E}[a|B,H]$) using the following lemma.

**Lemma 1** *Suppose $X$ and $Y$ are two jointly normal random variables with $X \sim \mathcal{N}(0,\sigma_X^2)$ and $Y \sim \mathcal{N}(0,\sigma_Y^2)$, and their correlation is $\rho$. For a given threshold $K$, we have the following conditional probability density functions (PDF):*

$$f_{X|Y\leq K}(x) = \frac{1}{\sqrt{2\pi}\sigma_X} exp\left(-\frac{x^2}{2\sigma_X^2}\right) \frac{\Phi\left(\frac{K-\rho\frac{\sigma_Y}{\sigma_X}x}{\sigma_Y\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{K}{\sigma_Y}\right)},$$

$$f_{X|Y>K}(x) = \frac{1}{\sqrt{2\pi}\sigma_X} exp\left(-\frac{x^2}{2\sigma_X^2}\right) \frac{\Phi\left(\frac{\rho\frac{\sigma_Y}{\sigma_X}x-K}{\sigma_Y\sqrt{1-\rho^2}}\right)}{\Phi\left(-\frac{K}{\sigma_Y}\right)},$$

*where $\Phi(\cdot)$ is the standard normal cumulative distribution function.*

We assume the threshold $K$ is set such that the number of High-risk tankers equals the number of Bad tankers, $\lambda$. Let $s^C$ denote the Clean exporters' optimally combined signal. Then, $\mathbb{E}(s^C \geq K) = \lambda$ implies

$$(1-\lambda)\Phi\left(\frac{K}{\sigma_s}\right) + \lambda\Phi\left(\frac{K-1}{\sigma_s}\right) = 1-\lambda,$$

where $\sigma_s$ is the noise standard deviation in signal $s^C$.

## F.2 Solving the model

(1) Substituting $\bar{w}_G^R$ and $\bar{w}_B^R$ into Rogue exporters' optimization problem gives

$$\begin{aligned}\max_{\{p_G^R, p_B^R\}} \quad & Q_{GL}^R(\tilde{r}^R - p_G^R - \bar{w}_{GL1}\cdot z^R) + Q_{GH}^R(\tilde{r}^R - p_G^R - \bar{w}_{GH1}\cdot z^R) + \\ & Q_{BL}^R(\tilde{r}^R - p_B^R - \bar{w}_{BL1}\cdot z^R) + Q_{BH}^R(\tilde{r}^R - p_B^R - \bar{w}_{BH1}\cdot z^R)\end{aligned}$$

Substituting market-clearing conditions into FOCs w.r.t. $p_G^R$ and $p_B^R$:

$$(1-\theta_G)(\tilde{r}^R - p_G^R - \bar{w}_{GL1}\cdot z^R - c_{GL}) + \theta_G(\tilde{r}^R - p_G^R - \bar{w}_{GH1}\cdot z^R - c_{GH}) = 0$$

$$(1-\theta_B)(\tilde{r}^R - p_B^R - \bar{w}_{BL1}\cdot z^R - c_{BL}) + \theta_B(\tilde{r}^R - p_B^R - \bar{w}_{BH1}\cdot z^R - c_{BH}) = 0$$

(2) Substituting $\bar{w}_L^C$ and $\bar{w}_H^C$ into Clean exporters' optimization problem gives

$$\begin{aligned}\max_{\{p_L,p_H\}} \quad & Q_{GL}(\tilde{r}-p_L-\bar{w}_{GL0}\cdot z^C)+Q_{GH}(\tilde{r}-p_H-\bar{w}_{GH0}\cdot z^C)+\\ & Q_{BL}(\tilde{r}-p_L-\bar{w}_{BL0}\cdot z^C)+Q_{BH}(\tilde{r}-p_H-\bar{w}_{BH0}\cdot z^C)\end{aligned}$$

Substituting market-clearing conditions into FOCs w.r.t. $p_L$ and $p_H$ and using the facts $\frac{\lambda(1-\theta_B)}{1-\lambda}=\theta_G$ and $\frac{(1-\lambda)\theta_G}{\lambda}=1-\theta_B$:

$$(1-\theta_G)(\tilde{r}-p_L-\bar{w}_{GL0}\cdot z^C+c_{GL}-\bar{c})+\theta_G(\tilde{r}-p_L-\bar{w}_{BL0}\cdot z^C+c_{BL}-\bar{c})=0$$

$$(1-\theta_B)(\tilde{r}-p_H-\bar{w}_{GL0}\cdot z^C+c_{GH}-\bar{c})+\theta_B(\tilde{r}-p_H-\bar{w}_{BH0}\cdot z^C+c_{BH}-\bar{c})=0$$

(3) Critical values for tanker decisions

$$\begin{aligned} c_{GL} &= p_G^R-p_L-(\bar{w}_{GL1}-\bar{w}_{GL0})z-\beta(V_G-V_B)\\ c_{GH} &= p_G^R-p_H-(\bar{w}_{GH1}-\bar{w}_{GH0})z-\beta(V_G-V_B)\\ c_{BL} &= p_B^R-p_L-(\bar{w}_{BL1}-\bar{w}_{BL0})z\\ c_{BH} &= p_B^R-p_H-(\bar{w}_{BH1}-\bar{w}_{BH0})z\end{aligned}$$

(4) Tanker values

$$V_G = \frac{\theta_G\mathbb{E}\left[\frac{(c_{GH})^2}{2\bar{c}}+p_H\right]+(1-\theta_G)\mathbb{E}\left[\frac{(c_{GL})^2}{2\bar{c}}+p_L\right]-[\theta_G w_{GH0}+(1-\theta_G)w_{GL0}]z}{1-\beta} \quad \text{(A-8)}$$

$$V_B = \frac{\theta_B\mathbb{E}\left[\frac{(c_{BH})^2}{2\bar{c}}+p_H\right]+(1-\theta_B)\mathbb{E}\left[\frac{(c_{BL})^2}{2\bar{c}}+p_L\right]-[\theta_B w_{BH0}+(1-\theta_B)w_{BL0}]z}{1-\beta} \quad \text{(A-9)}$$

and tankers' sanction penalty

$$z=\beta V_B$$

(5) Market clearing conditions

$$\begin{aligned} Q_{GL}^R &= (1-\lambda)(1-\theta_G)\left(\frac{c_{GL}}{\bar{c}}\right) & Q_{GL} &= (1-\lambda)(1-\theta_G)\left(1-\frac{c_{GL}}{\bar{c}}\right)\\ Q_{GH}^R &= (1-\lambda)\theta_G\left(\frac{c_{GH}}{\bar{c}}\right) & Q_{GH} &= (1-\lambda)\theta_G\left(1-\frac{c_{GH}}{\bar{c}}\right)\\ Q_{BL}^R &= \lambda(1-\theta_B)\left(\frac{c_{BL}}{\bar{c}}\right) & Q_{BL} &= \lambda(1-\theta_B)\left(1-\frac{c_{BL}}{\bar{c}}\right)\\ Q_{BH}^R &= \lambda\theta_B\left(\frac{c_{BH}}{\bar{c}}\right) & Q_{BH} &= \lambda\theta_B\left(1-\frac{c_{BH}}{\bar{c}}\right)\end{aligned}$$

(6) Clean exporters' information acquisition

$$\max_{\{\sigma_\xi,\chi\in\{0,1\}\}} \ \Pi(\sigma_\xi,\chi;\tilde{\sigma}_v^2)-\Omega(\sigma_\xi)-\chi\cdot P^W, \quad \text{(A-10)}$$

with optimality conditions

$$\frac{\partial \Pi(\sigma_\xi, \chi)}{\partial \sigma_\xi} = \frac{\partial \Omega(\sigma_\xi)}{\partial \sigma_\xi}, \quad \chi = \begin{cases} 0, & \text{if } \Pi(\sigma_\xi, 1) - \Pi(\sigma_\xi, 0) < P^W, \\ 1, & \text{if } \Pi(\sigma_\xi, 1) - \Pi(\sigma_\xi, 0) > P^W, \\ 0 \text{ or } 1, & \text{if } \Pi(\sigma_\xi, 1) - \Pi(\sigma_\xi, 0) = P^W. \end{cases} \tag{A-11}$$

There are 21 endogenous variables in the model: 4 prices ($p_L$, $p_H$, $p_G^R$, $p_B^R$), 4 critical values ($c_{GL}$, $c_{GH}$, $c_{BL}$, $c_{BH}$), 8 quantities ($Q_{GL}^R$, $Q_{GH}^R$, $Q_{BL}^R$, $Q_{BH}^R$, $Q_{GL}$, $Q_{GH}$, $Q_{BL}$, $Q_{BH}$), 2 tanker values ($V_G$, $V_B$), 1 sanction penalty $z$, and 2 variables about information acquisition ($\sigma_\xi$, $\chi$). We solve the model in the following manner.

First, for a given set of ($\sigma_\xi$, $\chi$), we solve the model by iterating on tanker values, $V_G$ and $V_B$. The steps are as follows:

- Step 1: Discretize $\tilde{\epsilon}$.
- Step 2: Set initial values as $V_G^{(0)}$ and $V_B^{(0)}$.
- Step 3: Given $V_G^{(k)}$ and $V_B^{(k)}$ from the $k$-th iteration, and for each discrete point of $\tilde{\epsilon}$, solve for the remaining endogenous variables (including 4 prices, 4 critical values, 8 quantities, and 1 sanction penalty) in a linear system, as those equations are linear.
- Step 4: Calculate $V_G^{(k+1,\text{raw})}$ and $V_B^{(k+1,\text{raw})}$ by substituting the values of the endogenous variables in Step 3 into Eq.(A-8) and (A-9). Update the tanker values:

  $$V_G^{(k+1)} = \phi V_G^{(k+1,\text{raw})} + (1-\phi) V_G^{(k)}, \quad V_B^{(k+1)} = \phi V_B^{(k+1,\text{raw})} + (1-\phi) V_B^{(k)},$$

  where $\phi$ is the learning rate.
- Step 5: Repeat Step 3 and 4 until convergence. The stopping condition is given by

  $$|V_G^{(k+1,\text{raw})} - V_G^{(k)}| + |V_B^{(k+1,\text{raw})} - V_B^{(k)}| < \eta,$$

  where $\eta$ is the tolerance.

We discretize $\tilde{\epsilon}$ on a grid of 101 points. We set the learning rate $\phi$ to 0.5 and tolerance $\eta$ to $10^{-7}$. The convergence of the above value iteration is ensured by the contraction mapping theorem. Our results are robust to different discretization parameters and learning rates.

Then, we deal with the Clean exporters' information-acquisition decision, $\sigma_\xi$ and $\chi$. In the pre-period equilibrium, the price of Windward's information, $P^W$, is determined such that Clean exporters are indifferent between buying information ($\chi = 1$) and not buying it ($\chi = 0$). Empirically, few Clean exporters purchased Windward's information in the pre-period. Accordingly, we approximate the pre-period equilibrium by $\chi = 0$. In the post-period, public disclosure sets $P^W = 0$ and induces $\chi = 1$. Given the choice of $\chi$, we solve the optimal choice of $\sigma_\xi$ by maximizing the Clean exporter's objective in Eq.(A-10).

## F.3 Details about calibrating $\tilde{\sigma}_v$

The system to solve is given by

$$x_i = \delta + \eta_i, \tag{A-12}$$

$$\tilde{\delta}_i = \mathbb{E}[\delta | x_i, M^W], \tag{A-13}$$

$$M^W = \sum_{i=1}^{N} \tilde{\delta}_i + \epsilon^W. \tag{A-14}$$

We conjecture the following linear solution:

$$\tilde{\delta}_i = \mu_\delta + A(x_i - \mu_\delta) + B\left(\frac{M^W}{N} - \mu_\delta\right), \tag{A-15}$$

where $A$ and $B$ are parameters to solve. Substituting Eq.(A-15) into Eq.(A-14) gives

$$\frac{M^W}{N} = \mu_\delta + \frac{A}{1-B}\left[(\delta + \bar{\eta} - \mu_\delta) + \frac{\epsilon^W}{AN}\right], \tag{A-16}$$

where $\bar{\eta} = \frac{1}{N}\sum_{i=1}^{N}\eta_i$. Define

$$X = (\delta + \bar{\eta} - \mu_\delta) + \frac{\epsilon^W}{AN}. \tag{A-17}$$

Note that $X$ is a sufficient statistic of $M^W$. Therefore, we have $\mathbb{E}[\delta|x_i, M^W] = \mathbb{E}[\delta|x_i, X]$. Applying the projection theorem for normal random variables gives

$$\begin{aligned}
\tilde{\delta}_i &= \mu_\delta + \begin{bmatrix} \mathrm{Cov}(x_i,\delta) & \mathrm{Cov}(X,\delta) \end{bmatrix} \begin{bmatrix} \mathrm{Var}(x_i) & \mathrm{Cov}(x_i,X) \\ \mathrm{Cov}(x_i,X) & \mathrm{Var}(X) \end{bmatrix}^{-1} \begin{bmatrix} x_i - \mu_\delta \\ X \end{bmatrix} \\
&= \mu_\delta + \begin{bmatrix} \Sigma_\delta & \Sigma_\delta \end{bmatrix} \begin{bmatrix} \Sigma_\delta + \Sigma_\eta & \Sigma_\delta + \frac{\Sigma_\eta}{N} \\ \Sigma_\delta + \frac{\Sigma_\eta}{N} & \Sigma_\delta + \frac{\Sigma_\eta}{N} + \frac{\Sigma_\epsilon}{A^2N^2} \end{bmatrix}^{-1} \begin{bmatrix} x_i - \mu_\delta \\ X \end{bmatrix} \\
&= \mu_\delta + \frac{\frac{1}{A^2N^2}\Sigma_\epsilon\Sigma_\delta(x_i - \mu_\delta) + \frac{N-1}{N}\Sigma_\eta\Sigma_\delta X}{(\Sigma_\delta + \Sigma_\eta)\left(\Sigma_\delta + \frac{\Sigma_\eta}{N} + \frac{\Sigma_\epsilon}{A^2N^2}\right) - \left(\Sigma_\delta + \frac{\Sigma_\eta}{N}\right)^2}.
\end{aligned} \tag{A-18}$$

On the other hand, combining Eq.(A-15) and Eq.(A-17) gives

$$\tilde{\delta}_i = \mu_\delta + A(x_i - \mu_\delta) + \frac{AB}{1-B}X. \tag{A-19}$$

Comparing coefficients in Eq.(A-18) and Eq.(A-19) yields

$$\frac{\Sigma_\eta}{\Sigma_\epsilon}(N-1)\left(N + \frac{\Sigma_\eta}{\Sigma_\delta}\right)A^3 + A - 1 = 0, \tag{A-20}$$

$$\frac{B}{1-B} = N(N-1)\frac{\Sigma_\eta}{\Sigma_\epsilon}A^2. \tag{A-21}$$

Finally, we prove that when $\Sigma_\delta/\Sigma_\eta \to \underline{c} > 0$ and $\Sigma_\epsilon/\Sigma_\eta \to 0$, we have $A \to 0$ and $B \to 1$. Note that Eq.(A-20) implies that

$$0 < A < \left(\frac{\Sigma_\epsilon}{\Sigma_\eta}\right)^{\frac{1}{3}} (N-1)^{-\frac{1}{3}} N^{-\frac{1}{3}}. \tag{A-22}$$

Therefore, when $\Sigma_\epsilon/\Sigma_\eta \to 0$, we have $A \to 0$. Note that Eq.(A-20) and Eq.(A-21) imply

$$\frac{B}{1-B} = N(N-1)\frac{\Sigma_\eta}{\Sigma_\epsilon}A^2 = \frac{A^{-1}-1}{1+\frac{1}{N}\frac{\Sigma_\eta}{\Sigma_\delta}}. \tag{A-23}$$

Therefore, when $A \to 0$ and $1+\frac{1}{N}\frac{\Sigma_\eta}{\Sigma_\delta} \to 1+\frac{1}{N\bar{c}}$, we have $\frac{B}{1-B} \to \infty$ and $B \to 1$.

## F.4 Upper bound of the valuation of Windward's signal

Let $M_{signal}$ and $M_{other}$ denote the market values of Windward's signal and its other businesses, respectively. Let superscripts 0 and 1 refer to the pre- and post-periods, respectively. Thus, $M^0_{signal}$ is the pre-period value of Windward's signal. The data imply

$$M^0_{signal} + M^0_{other} = \$42\text{m}.$$

Also, Windward's market value increases in the post-period

$$M^1_{signal} + M^1_{other} \geq M^0_{signal} + M^0_{other}.$$

Public disclosure eliminates the proprietary value of the signal

$$M^1_{signal} = 0.$$

Assume the post-period spillover to other businesses (i.e., the market updates $M^0_{signal}$ to $M^1_{signal}$ also based on the signal precision) is bounded by $b$:

$$\frac{M^1_{other}}{M^0_{other}} \leq b.$$

Therefore, we have

$$\begin{aligned} M^0_{signal} &= \left(1 - \frac{M^0_{other}}{M^0_{signal} + M^0_{other}}\right) \times \$42\text{m} \leq \left(1 - \frac{M^0_{other}}{M^1_{signal} + M^1_{other}}\right) \times \$42\text{m} \\ &= \left(1 - \frac{M^0_{other}}{M^1_{other}}\right) \times \$42\text{m} \leq \left(1 - \frac{1}{b}\right) \times \$42\text{m} \end{aligned}$$

If $b$ is tied to the change in signal precision—from 7.5% pre to 21.3% post—then $b = 2.84$, yielding

$$M^0_{signal} \leq \$27\text{m}.$$

## F.5 Model analogs

- First, the number of sanctioned tankers in the model is

$$Q_{Sanc} = \sum_{ij \in \{GL,GH,BL,BH\}} \left( Q_{ij} \bar{w}_{ij0} + Q_{ij}^{R} \bar{w}_{ij1} \right). \tag{A-24}$$

  Since we normalize the total number of tankers to one, $Q_{Sanc}$ is also the proportion of sanctioned tankers.

- Second, the volatilities of low-risk and high-risk tankers' fixtures are std($p_L$) and std($p_H$), respectively.

- Third, we derive the model-implied pseudo-$R^2$ for predicting to-be-sanctioned tankers. Generally, let $y = 0, 1$ denote the true label and $\hat{y} = 0, 1$ denote the prediction. Pseudo-$R^2$ is a function of three elements: proportion of sanctioned tankers $\bar{p} = \mathbb{E}[y]$; false positive rate $\alpha = \mathbb{E}[\hat{y}|y = 0]$; and true positive rate $\beta = \mathbb{E}[\hat{y}|y = 1]$. Define $\hat{p} = \mathbb{E}[y|\hat{y}]$, which gives $\hat{p}_0 = \mathbb{E}[y|\hat{y} = 0] = \frac{\bar{p}(1-\beta)}{(1-\bar{p})(1-\alpha)+\bar{p}(1-\beta)}$ and $\hat{p}_1 = \mathbb{E}[y|\hat{y} = 1] = \frac{\bar{p}\beta}{(1-\bar{p})\alpha+\bar{p}\beta}$. Pseudo-$R^2$ equals

$$\begin{aligned} & \text{pseudo-}R^2 \\ = \; & 1 - \frac{\mathbb{E}[y\log(\hat{p}) + (1-y)\log(1-\hat{p})]}{\mathbb{E}[y\log(\bar{p}) + (1-y)\log(1-\bar{p})]} \\ = \; & 1 - \frac{\bar{p}(1-\beta)\log(\hat{p}_0) + \bar{p}\beta\log(\hat{p}_1) + (1-\bar{p})(1-\alpha)\log(1-\hat{p}_0) + (1-\bar{p})\alpha\log(1-\hat{p}_1)}{\bar{p}\log(\bar{p}) + (1-\bar{p})\log(1-\bar{p})}. \end{aligned} \tag{A-25}$$

  When $y = 1$ ($y = 0$) indicates a tanker being sanctioned (not being sanctioned), and $\hat{y} = 1$ ($\hat{y} = 0$) indicates a tanker labeled as high-risk (low-risk), we have the following classification:

| | Low-risk ($\hat{y} = 0$) | High-risk ($\hat{y} = 1$) |
|---|---|---|
| Non-sanctioned ($y = 0$) | $(1-\bar{w}_{BL}) \cdot \text{BL} + (1-\bar{w}_{GL}) \cdot \text{GL}$ | $(1-\bar{w}_{BH}) \cdot \text{BH} + (1-\bar{w}_{GH}) \cdot \text{GH}$ |
| Sanctioned ($y = 1$) | $\bar{w}_{BL} \cdot \text{BL} + \bar{w}_{GL} \cdot \text{GL}$ | $\bar{w}_{BH} \cdot \text{BH} + \bar{w}_{GH} \cdot \text{GH}$ |

  GL, GH, BL, and BH are proportions of tankers in each category; $\bar{w}_{GL}$, $\bar{w}_{GH}$, $\bar{w}_{BL}$, and $\bar{w}_{BH}$ are equilibrium sanctioned probabilities conditional on each category. In equilibrium, we have

$$\bar{w}_{xy} = \bar{w}_{xy0} + \frac{c_{xy}}{\bar{c}}(\bar{w}_{xy1} - \bar{w}_{xy0}), \quad xy \in \{GL, GH, BL, BH\}$$

  where $\bar{w}_{xy0}$ and $\bar{w}_{xy1}$ are defined in the paper, and $\frac{c_{xy}}{\bar{c}}$ captures the probabilities of dealing with Rogue exporters in the current period. Therefore, based on the table, we have

$$\begin{aligned} \bar{p} &= \bar{w}_{BL} \cdot \text{BL} + \bar{w}_{GL} \cdot \text{GL} + \bar{w}_{BH} \cdot \text{BH} + \bar{w}_{GH} \cdot \text{GH}, \\ \alpha &= \frac{(1-\bar{w}_{BH}) \cdot \text{BH} + (1-\bar{w}_{GH}) \cdot \text{GH}}{(1-\bar{w}_{BL}) \cdot \text{BL} + (1-\bar{w}_{GL}) \cdot \text{GL} + (1-\bar{w}_{BH}) \cdot \text{BH} + (1-\bar{w}_{GH}) \cdot \text{GH}}, \end{aligned}$$

$$\beta = \frac{\bar{w}_{BH} \cdot \mathrm{BH} + \bar{w}_{GH} \cdot \mathrm{GH}}{\bar{w}_{BL} \cdot \mathrm{BL} + \bar{w}_{GL} \cdot \mathrm{GL} + \bar{w}_{BH} \cdot \mathrm{BH} + \bar{w}_{GH} \cdot \mathrm{GH}}.$$

Substituting them into Eq.(A-25) gives the pseudo-$R^2$ for predicting to-be-sanctioned tankers.

- Then, the model analog for the disclosure effect on high-risk tankers' fixtures is

$$\frac{\mathbb{E}[p^{post} - p^{pre} | \mathrm{H}^{post}]}{\mathbb{E}[p^{pre} | \mathrm{H}^{post}]}, \tag{A-26}$$

where $p$ is fixture and $\mathrm{H}^{post}$ denotes a tanker that is labeled as high-risk in the post-period. The average pre-period fixture, conditional on a post-period high-risk label is

$$\mathbb{E}[p^{pre} | \mathrm{H}^{post}] = \mathbb{P}[\mathrm{L}^{pre} | \mathrm{H}^{post}] \times \mathbb{E}[p_L^{pre}] + \mathbb{P}[\mathrm{H}^{pre} | \mathrm{H}^{post}] \times \mathbb{E}[p_H^{pre}], \tag{A-27}$$

where $\mathbb{E}[p_L^{pre}]$ and $\mathbb{E}[p_H^{pre}]$ are price averages taken over the pre-period aggregate shock $\tilde{\epsilon}$ and $\mathbb{P}[.|.]$ is conditional probability. Noting that $\mathbb{P}[\mathrm{L}^{pre} | \mathrm{H}^{post}] = 1 - \mathbb{P}[\mathrm{H}^{pre} | \mathrm{H}^{post}]$, we calculate $\mathbb{P}[\mathrm{H}^{pre} | \mathrm{H}^{post}]$ by splitting it into three cases based on tanker types: (i) $\mathrm{G}^{pre}$ and $\mathrm{G}^{post}$, (ii) $\mathrm{G}^{pre}$ and $\mathrm{B}^{post}$, and (iii) $\mathrm{B}^{pre}$ and $\mathrm{B}^{post}$. This is detailed below.

We use subscripts 0 and 1 to denote the pre- and post-period, respectively. For example, $H_0$ denotes the event of a tanker being labeled as high-risk in the pre-period. Using the chain rule of conditional probabilities, we have

$$\begin{aligned}
\mathbb{P}[H_0|H_1] &= \mathbb{P}[H_0|G_1, H_1] \times \mathbb{P}[G_1|H_1] + \mathbb{P}[H_0|B_1, H_1] \times \mathbb{P}[B_1|H_1] \\
&= \mathbb{P}[H_0|G_0, G_1, H_1] \times \mathbb{P}[G_0|G_1, H_1] \times \mathbb{P}[G_1|H_1] + \\
&\quad \mathbb{P}[H_0|B_0, G_1, H_1] \times \mathbb{P}[B_0|G_1, H_1] \times \mathbb{P}[G_1|H_1] + \\
&\quad \mathbb{P}[H_0|G_0, B_1, H_1] \times \mathbb{P}[G_0|B_1, H_1] \times \mathbb{P}[B_1|H_1] + \\
&\quad \mathbb{P}[H_0|B_0, B_1, H_1] \times \mathbb{P}[B_0|B_1, H_1] \times \mathbb{P}[B_1|H_1]
\end{aligned}$$

We calculate each term as follows. First, as per Section 6.1.1 in the paper, we have $\mathbb{P}[G_1|H_1] = 1 - \theta_{B,1}$ and $\mathbb{P}[B_1|H_1] = \theta_{B,1}$. Second, because in the model risk labels are generated based on current tanker types only (i.e., the classification technology has i.i.d. errors), the post-period tanker type is a sufficient statistic for the pre-period tanker type. That is:

$$\begin{aligned}
\mathbb{P}[G_0|G_1, H_1] &= \mathbb{P}[G_0|G_1], \qquad \mathbb{P}[B_0|G_1, H_1] = \mathbb{P}[B_0|G_1], \\
\mathbb{P}[G_0|B_1, H_1] &= \mathbb{P}[G_0|B_1], \qquad \mathbb{P}[B_0|B_1, H_1] = \mathbb{P}[B_0|B_1].
\end{aligned}$$

Because the change of a G tanker to a B tanker is irreversible, we have $\mathbb{P}[G_0|G_1] = 1$ and $\mathbb{P}[B_0|G_1] = 0$. Since $B_1$ tankers come from (i) $B_0$ tankers that have not been sanctioned or (ii) $G_0$ tankers that have dealt with Rogue exporters in period 0 and have not been

sanctioned, we have

$$\begin{aligned}\mathbb{P}[B_0|B_1] &= \frac{\tilde{Q}_{BL,0}+\tilde{Q}_{BH,0}+\tilde{Q}^R_{BL,0}+\tilde{Q}^R_{BH,0}}{\tilde{Q}_{BL,0}+\tilde{Q}_{BH,0}+\tilde{Q}^R_{BL,0}+\tilde{Q}^R_{BH,0}+\tilde{Q}^R_{GL,0}+\tilde{Q}^R_{GH,0}},\\ \mathbb{P}[G_0|B_1] &= \frac{\tilde{Q}^R_{GL,0}+\tilde{Q}^R_{GH,0}}{\tilde{Q}_{BL,0}+\tilde{Q}_{BH,0}+\tilde{Q}^R_{BL,0}+\tilde{Q}^R_{BH,0}+\tilde{Q}^R_{GL,0}+\tilde{Q}^R_{GH,0}},\end{aligned}$$

where $\tilde{Q}_{BL,0}$, $\tilde{Q}_{BH,0}$, $\tilde{Q}^R_{BL,0}$, $\tilde{Q}^R_{BH,0}$, $\tilde{Q}^R_{GL,0}$, $\tilde{Q}^R_{GH,0}$ are the number of tankers for each type surviving sanction in pre-period:

$$\begin{aligned}\tilde{Q}_{BL,0}=Q_{BL,0}(1-w_{L,0}), &\qquad \tilde{Q}_{BH,0}=Q_{BH,0}(1-w_{H,0}),\\ \tilde{Q}^R_{BL,0}=Q^R_{BL,0}(1-w_{L,0}-\delta_{L,0}), &\qquad \tilde{Q}^R_{BH,0}=Q^R_{BH,0}(1-w_{H,0}-\delta_{H,0}),\\ \tilde{Q}^R_{GL,0}=Q^R_{GL,0}(1-\delta_{L,0}), &\qquad \tilde{Q}^R_{GH,0}=Q^R_{GH,0}(1-\delta_{H,0}).\end{aligned}$$

Finally, again because risk labels only depend on current tanker types, we have

$$\begin{aligned}\mathbb{P}[H_0|G_0,G_1,H_1]=\mathbb{P}[H_0|G_0]=\theta_{G,0}, &\qquad \mathbb{P}[H_0|B_0,G_1,H_1]=\mathbb{P}[H_0|B_0]=\theta_{B,0},\\ \mathbb{P}[H_0|G_0,B_1,H_1]=\mathbb{P}[H_0|G_0]=\theta_{G,0}, &\qquad \mathbb{P}[H_0|B_0,B_1,H_1]=\mathbb{P}[H_0|B_0]=\theta_{B,0}.\end{aligned}$$

Combining the above equations yields

$$\begin{aligned}\mathbb{P}[H_0|H_1] &= \theta_{G,0}(1-\theta_{B,1})+\\ &\quad \theta_{G,0}\theta_{B,1}\left(\frac{\tilde{Q}^R_{GL,0}+\tilde{Q}^R_{GH,0}}{\tilde{Q}_{BL,0}+\tilde{Q}_{BH,0}+\tilde{Q}^R_{BL,0}+\tilde{Q}^R_{BH,0}+\tilde{Q}^R_{GL,0}+\tilde{Q}^R_{GH,0}}\right)+\\ &\quad \theta_{B,0}\theta_{B,1}\left(\frac{\tilde{Q}_{BL,0}+\tilde{Q}_{BH,0}+\tilde{Q}^R_{BL,0}+\tilde{Q}^R_{BH,0}}{\tilde{Q}_{BL,0}+\tilde{Q}_{BH,0}+\tilde{Q}^R_{BL,0}+\tilde{Q}^R_{BH,0}+\tilde{Q}^R_{GL,0}+\tilde{Q}^R_{GH,0}}\right).\end{aligned}$$

- Finally, we use the probability of sending AIS signals from areas near sanctioned countries as a proxy for the probability of dealing with Rogue exporters in the model. Defining the probabilities for H and L tankers of dealing with Rogue exporters as:

$$\begin{aligned}b_H &= \mathbb{P}[\mathrm{B}|\mathrm{H}]\times\mathbb{P}[\tilde{c}_i<c_{BH}]+\mathbb{P}[\mathrm{G}|\mathrm{H}]\times\mathbb{P}[\tilde{c}_i<c_{GH}]=\theta_B\frac{c_{BH}}{\bar{c}}+(1-\theta_B)\frac{c_{GH}}{\bar{c}},\\ b_L &= \mathbb{P}[\mathrm{B}|\mathrm{L}]\times\mathbb{P}[\tilde{c}_i<c_{BL}]+\mathbb{P}[\mathrm{G}|\mathrm{L}]\times\mathbb{P}[\tilde{c}_i<c_{GL}]=\theta_G\frac{c_{BL}}{\bar{c}}+(1-\theta_G)\frac{c_{GL}}{\bar{c}},\end{aligned}$$

the disclosure's effect on high-risk tankers' routes in the model is

$$\frac{\mathbb{E}[b^{post}-b^{pre}|\mathrm{H}^{post}]}{\mathbb{E}[b^{pre}|\mathrm{H}^{post}]}, \tag{A-28}$$

where

$$\mathbb{E}[b^{pre}|\mathrm{H}^{post}]=\mathbb{P}[\mathrm{L}^{pre}|\mathrm{H}^{post}]\times\mathbb{E}[b^{pre}_L]+\mathbb{P}[\mathrm{H}^{pre}|\mathrm{H}^{post}]\times\mathbb{E}[b^{pre}_H]. \tag{A-29}$$

$\mathbb{E}[b^{pre}_L]$ and $\mathbb{E}[b^{pre}_H]$ are the averages of the respective probabilities taken over the pre-period aggregate shock $\tilde{\epsilon}$ and the conditional probability $\mathbb{P}[\mathrm{H}^{pre}|\mathrm{H}^{post}]$ is as specified above.

## F.6 Model-implied benefits for exporters

Table A-8 shows the model-implied benefits from the disclosure for the top 20 exporters. The gain for each one is greater than the Windward's stock market valuation, $42 mln.

Table A-8: Model-implied benefits for top 20 exporters

This table shows the model-implied benefits from the disclosure for the top 20 exporters. We calculate the market share using the number of tanker-trips for each exporter divided by the total number of tanker-trips during one year before disclosure, i.e., August 2022 - July 2023. The dollar benefit equals the model-implied total benefit, $4,299 mln, multiplied by the market share.

| Rank | Exporter | Share | Benefits ($ mln) | Rank | Exporter | Share | Benefits ($ mln) |
|---|---|---|---|---|---|---|---|
| 1 | BP | 6.16% | 264.8 | 11 | LITASCO | 2.46% | 105.8 |
| 2 | VITOL | 4.78% | 205.6 | 12 | ST SHIPPING | 2.42% | 104.2 |
| 3 | SHELL | 4.73% | 203.4 | 13 | ARAMCO | 2.42% | 104.2 |
| 4 | REPSOL | 4.40% | 189.2 | 14 | PETROBRAS | 2.42% | 104.2 |
| 5 | EXXON MOBIL | 3.92% | 168.3 | 15 | TOTAL | 2.33% | 100.3 |
| 6 | UNIPEC | 3.34% | 143.6 | 16 | ENI | 2.19% | 94.3 |
| 7 | CSSSA | 2.87% | 123.4 | 17 | CSSA | 1.75% | 75.1 |
| 8 | TRAFIGURA | 2.69% | 115.7 | 18 | AMPOL | 1.66% | 71.3 |
| 9 | CNR | 2.69% | 115.7 | 19 | ADMIC | 1.62% | 69.6 |
| 10 | CHEVRON | 2.47% | 106.4 | 20 | CEPSA | 1.59% | 68.5 |

# References


Abadie, A. (2005), 'Semiparametric Difference-in-Differences Estimators', *Review of Economic Studies* **72**(1), 1–19.

Bai, X., Fernández-Villaverde, J., Li, Y., Xu, L. and Zanetti, F. (2025), 'The (Un)Intended Consequences of Oil Sanctions Through the Dark Shipping of Sanctioned Oil', *Working Paper* .

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997), 'Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme', *Review of Economic Studies* **64**(4), 605–654.

OFAC (2020), Sanctions Advisory for the Maritime Industry, Energy and Metal Sectors, and Related Communities, U.S. Department of the Treasury, Office of Foreign Assets Control.

Pallotta, G., Vespe, M. and Bryan, K. (2013), 'Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction', *Entropy* **15**(6), 2218–2245.

Rosenbaum, P. R. and Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Weng, J., Li, G. and Zhao, Y. (2022), 'Detection of abnormal ship trajectory based on the complex polygon', *Journal of Navigation* **75**(4), 966–983.