Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Why the G7 Oil Price Cap Failed: Shadow Fleets,War Premium Decay, and Lessons for Sanctions Design

Tuğba Ortay Baykal

tortay.gokce@istanbulticaret.edu.tr

İstanbul Ticaret Üniversitesi: Istanbul Ticaret Universitesi https://orcid.org/0009-0001-1414-2523

Research Article

Keywords: Oil sanctions, Price cap, Russia, Shadow fleet, Sanctions evasion, Energy security

Posted Date: March 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8881992/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Why the G7 Oil Price Cap Failed: Shadow Fleets, War Premium Decay, and Lessons for Sanctions Design

Tuğba Ortay Baykal[1*]

[1*]Department of Economics, Istanbul Ticaret University, Istanbul, Turkey.

Corresponding author(s). E-mail(s): tugbaortaygokce@gmail.com;

Abstract

The G7/EU oil price cap on Russian crude, implemented in December 2022 at $60 per barrel, was designed to reduce Russian revenues while maintaining global supply. Using daily price data and 25,399 vessel-level cargo records (January 2020–January 2026), we find the price cap did not function as a binding constraint. The war outbreak is associated with the largest change in the Brent-Urals spread (cumulative abnormal returns of $22.35/bbl, $\boldsymbol{p < 0.001}$), with no significant market reactions to subsequent sanction announcements. The apparent spread narrowing under the cap becomes statistically insignificant when measured as a percentage of Brent, and is absorbed entirely by a war premium decay model (half-life 2.4 years). Russia's shadow fleet grew from 40.5% (pre-war period average) to 62.6% (2024 annual peak) of seaborne exports, yet targeted vessel sanctions (January 2025) proved more effective than the price cap, reducing shadow fleet share to 51.7% (period average). A Pacific versus Atlantic quasi-difference-in-differences analysis suggests that spread narrowing was not driven by the EU embargo alone. Cumulative post-war revenue losses are estimated at $87–130 billion (central estimate $109 billion, based on spot spreads and seaborne volumes, excluding pipeline exports), of which shadow fleet expansion recovered only $2.2 billion. Our identification relies on event studies and conditional associations rather than quasi-experimental variation; nonetheless, the findings suggest that policymakers should prioritize targeted vessel designations and flag-state coordination over price-level adjustments.

Keywords: Oil sanctions, Price cap, Russia, Shadow fleet, Sanctions evasion, Energy security

JEL Classification: F51 , Q41 , Q43 , G14

1 Introduction

On February 24, 2022, Russia launched a full-scale invasion of Ukraine, triggering the most extensive sanctions regime in modern history. Among the most significant measures was the G7/EU price cap on Russian crude oil, implemented on December 5, 2022, at $60 per barrel. The cap was designed to reduce Russian oil revenues while maintaining global supply by allowing Russian oil to reach world markets using Western shipping services only if priced below the threshold (Johnson et al 2023). Unlike traditional embargoes, the cap sought to exploit the West's dominance in maritime insurance and shipping to constrain the price Russia could receive while preserving incentives for continued production (Wachtmeister et al 2023).

However, the effectiveness of this novel mechanism has been questioned. Russia rapidly expanded its "shadow fleet" of tankers operating outside Western insurance and service networks, potentially circumventing the cap entirely (Cardoso et al 2025). The broader sanctions literature suggests that economic sanctions succeed in roughly one-third of cases, with effectiveness declining when targets are large economies (Hufbauer et al 2007). Felbermayr et al (2024) show that modern sanctions increasingly target financial and trade networks, and that enforcement and evasion dynamics are central to outcomes. Itskhoki and Ribakova (2024) highlight the gap between intended and realized effects in the Russian case specifically.

The price cap's theoretical foundations have received substantial attention. Johnson et al (2023) explain that enforcement operates through maritime services control rather than physical interdiction. de Souza et al (2024) formalize trade sanction design, showing that optimal policy depends on the sanctioner's ability to control intermediary networks. Wachtmeister et al (2023) use field-level Russian production cost data to show that the optimal cap lies between $40–60/barrel, suggesting the actual cap was set conservatively. Turner and Sappington (2024) demonstrate that a price cap can paradoxically increase Russia's output and revenues.

Empirically, several studies have assessed the cap's effectiveness. Spiro et al (2025) synthesize the evidence, concluding that the EU import embargo was more consequential than the cap, which functioned as a "partial transport embargo" rather than a binding price constraint. Kilian et al (2024) decompose the Russian oil discount into a buyer's bargaining power component and a transport cost component, with the embargo's market-segmentation effect dominating. Babina et al (2025) construct a transaction-level dataset from Russian customs data and document systematic cap violations, particularly for ESPO crude exported from Pacific ports. Cardoso et al (2025) show theoretically that the cap's primary effect was to incentivize shadow fleet expansion rather than bind prices.

The shadow fleet itself has attracted growing research attention. Fernandez-Villaverde et al (2025) estimate the dark fleet averaged 558 tankers during 2017–2023, comprising roughly one-quarter of the global crude tanker fleet. Rodriguez-Diaz et al (2025) distinguish "dark fleets" (wholly illicit) from "gray fleets" (exploiting legal ambiguities). Levi et al (2023) map shadow tanker operations empirically, documenting ownership structures and trade routes. In the event study literature, Omar and Lambe (2022) find that sanctions on oil exporters raise prices, imposing costs on the

sanctioner, and Kokeny and Kelemen (2025) show that only surprise events generate significant returns in the Russia-Ukraine context.

Our study contributes to this literature in ways that complement and extend the qualitative assessment of Spiro et al (2025). While their analysis relies on descriptive decomposition and heuristic reasoning, we employ formal econometric identification: event study methodology, level and first-differenced regressions with cointegration analysis, placebo tests, VAR modeling, structural break analysis, and a Pacific-Atlantic quasi-difference-in-differences design. Four features distinguish our approach. First, we construct a vessel-level shadow fleet database from 25,399 cargo records across 1,636 vessels, merging LSEG World Flows data with FleetLeaks sanctioned vessel lists from six jurisdictions, enabling precise tracking of shadow fleet market share evolution that is unavailable in aggregate price studies. Second, we document a war premium decay model whose half-life (2.4 years exponential, 2.8 years linear) absorbs both the price cap effect and the shadow fleet effect, a finding not reported in prior work. Third, we exploit the Pacific (ESPO) versus Atlantic (Urals) route distinction in a quasi-difference-in-differences framework that separates the EU embargo from the price cap mechanism. Fourth, our sample extends through January 2026, uniquely capturing the January 2025 Biden tanker sanctions and their measurable 11.4 percentage point reduction in shadow fleet share.

This study addresses three questions: (1) How did major sanction announcements affect the Brent-Urals spread? (2) Did the price cap function as a binding constraint? (3) Has shadow fleet expansion weakened sanction effectiveness? Our central finding is that the cap did not function as intended, with implications for the ongoing policy debate on cap-level adjustments versus targeted enforcement.

The remainder of this paper is organized as follows. Section 2 describes our data and methodology. Section 3 presents results and discussion. Section 4 offers conclusions and policy implications.

2 Data and Methodology

Price Data. Daily oil price data are obtained from Refinitiv Eikon for January 2, 2020 to January 24, 2026 ($N = 1{,}583$ trading days). Variables include Brent Crude (FOB), Urals Baltic (CIF), Urals Novorossiysk (CIF), ESPO (FOB Kozmino), and BDTI (Baltic Dirty Tanker Index). We calculate the spread as: Spread = Brent$-$Urals Baltic. A caveat: Brent is FOB while Urals is CIF, so the spread conflates the pure discount with transport costs. We address this by including BDTI as a control and constructing a freight-adjusted spread (Appendix B). ESPO crude from Russia's Pacific ports is included as a robustness measure; unlike Urals, ESPO is priced FOB and traded predominantly to China, with 91.6% of post-cap observations priced above the $60 threshold, confirming widespread cap violations on the Pacific route documented by Babina et al (2025).

Controls. US Dollar Index (DXY) and CBOE Volatility Index (VIX).[1]

[1] S&P 500 returns were also obtained but excluded from regressions due to high collinearity with VIX ($r = -0.72$).

Cargo Data. From LSEG World Flows covering January 2019 to February 2026: 221,102 global cargo records, of which 25,399 depart from Russian ports aboard 1,636 unique vessels.

Shadow Fleet Identification. We use FleetLeaks sanctioned vessel databases from six jurisdictions (US/OFAC, EU, UK, Canada, Australia, New Zealand). The "ALL" database contains 882 vessels; the "US" database contains 167 OFAC-designated vessels.

Our primary shadow fleet definition (Definition B) classifies a vessel as shadow fleet if it is either (i) sanctioned by at least one jurisdiction, or (ii) aged $\geq$20 years *and* registered under a convenience flag (Panama, Liberia, Marshall Islands, Palau, Cameroon, Gabon, and others per the ITF classification). The age-flag criterion captures the economic logic of sanctions evasion: vessels approaching scrapping age have low residual market value, making them available for purchase by entities willing to accept sanctions risk, while convenience flags offer reduced regulatory oversight and allow opaque ownership structures that obscure beneficial control (Rodriguez-Diaz et al 2025; Fernandez-Villaverde et al 2025). We also report results using Definition A (narrow: sanctioned only) and Definition C (broad: age $\geq$15 years). Main analysis uses Definition B throughout.

Sample Characteristics. Table 1 presents summary statistics by period. The Brent-Urals spread averaged \$1.61/bbl before the war, increased to \$29.91/bbl in the war pre-cap period, then narrowed to \$17.43/bbl post-cap and \$12.86/bbl post-Biden. Shadow fleet share rose from 40.5% pre-war to 63.1% post-cap before declining to 51.7% post-Biden.

Table 1 Summary Statistics by Period

Variable	Full Sample	Pre-War	War Pre-Cap	Post-Cap	Post-Biden
Panel A: Prices ($/bbl)					
Brent Crude	74.76	58.91	105.80	82.18	69.41
	(20.14)	(19.43)	(12.42)	(5.83)	(5.28)
Urals Baltic	62.12	57.30	75.89	64.74	56.55
	(14.45)	(18.77)	(9.72)	(8.39)	(6.69)
Brent-Urals Spread	12.64	1.61	29.91	17.43	12.86
	(10.47)	(1.69)	(5.40)	(6.44)	(2.78)
Panel B: Volumes					
Daily Volume (KT)	748.66	697.13	770.56	787.21	760.93
	(234.14)	(228.92)	(212.91)	(239.98)	(229.88)
Shadow Fleet Share (%)	50.7	40.5	43.8	63.1	51.7
	(20.2)	(17.9)	(16.7)	(17.3)	(18.9)
Panel C: Controls					
BDTI	1,017	685	1,470	1,150	1,099
DXY	100.11	94.28	105.48	103.91	100.48
VIX	20.88	24.34	26.25	16.41	18.80
Observations	1,583	561	202	549	271

Notes: Standard deviations in parentheses. Pre-War: Jan 2020–Feb 2022. War Pre-Cap: Feb–Dec 2022. Post-Cap: Dec 2022–Jan 2025. Post-Biden: Jan 2025+.

Event Study. We employ standard event study methodology following MacKinlay (1997) and Kokeny and Kelemen (2025). For each of six key sanction-related events, we define an estimation window $[-120, -6]$ and event window $[-5, +10]$. We estimate the market model:

$$R_t = \alpha + \beta R_{m,t} + \epsilon_t \tag{1}$$

where R_t is the daily change in the Brent-Urals spread and $R_{m,t}$ is the daily return on Brent crude. Abnormal returns and cumulative abnormal returns (CAR) are computed as:

$$AR_t = R_t - (\hat{\alpha} + \hat{\beta} R_{m,t}), \qquad CAR[-5, +10] = \sum_{t=-5}^{+10} AR_t \tag{2}$$

Regression Model. We estimate:

$$\text{Spread}_t = \alpha + \beta_1 \text{PostWar}_t + \beta_2 \text{PostCap}_t + \beta_3 \text{PostBiden}_t + \gamma \text{Shadow}_t + \boldsymbol{\delta}' \mathbf{X}_t + \epsilon_t \tag{3}$$

where $\text{PostWar}_t = \mathbf{1}(t > \text{Feb 24, 2022})$, $\text{PostCap}_t = \mathbf{1}(t > \text{Dec 5, 2022})$, $\text{PostBiden}_t = \mathbf{1}(t > \text{Jan 10, 2025})$, and $\mathbf{X}_t$ includes DXY, VIX, BDTI, and a COVID crash dummy (equal to one during March–June 2020, capturing the demand shock and negative WTI episode). The step dummies are cumulative: $\text{PostWar}_t = 1$ for all observations after February 24, 2022, including the post-cap and post-Biden subperiods. Thus β_2 captures the *incremental* association of the cap period relative to the war-only period. Standard errors use the robust heteroskedasticity and autocorrelation consistent estimator of Newey and West (1987) with 10 lags.

Identification and Interpretation. An important caveat is warranted. The post-event step dummies capture the combined effect of all factors that changed after each date, not solely sanctions. Between February 2022 and December 2022, global oil supply/demand fundamentals shifted, EU embargo announcements created anticipation effects, and OPEC+ production decisions evolved. Our controls absorb some confounders but cannot fully isolate sanctions effects from concurrent macroeconomic changes. We therefore interpret level regression coefficients as conditional associations, not as causal effects of sanctions per se.

A particularly important confound concerns the PostCap dummy. The G7 price cap took effect on December 5, 2022, the same week that the EU's physical embargo on seaborne Russian crude imports became operational. Our regression cannot fully separate the price mechanism (the cap) from the quantity mechanism (the embargo). Kilian et al (2024) decompose the Russian oil discount and find that the embargo's market-segmentation effect dominates. To address this confound, we exploit the Pacific (ESPO) versus Atlantic (Urals) asymmetry: the EU embargo restricts only Atlantic-route imports, while the price cap applies to both routes. If the embargo drove the post-cap narrowing, the Atlantic spread should narrow more than the Pacific spread. Our quasi-difference-in-differences test finds no significant difference (Section 3). Similarly, the PostWar dummy absorbs not only the invasion itself but also voluntary self-sanctions by Western buyers, EU sanctions package discussions, and OPEC+ supply management decisions.

The event study, which exploits narrow time windows, provides stronger identification for immediate market reactions, though post-war events face the additional challenge that the estimation window $[-120, -6]$ spans a period of structural change.

Time Series Properties. The Brent-Urals spread is non-stationary in levels (ADF $p = 0.26$; KPSS $p < 0.01$) but stationary in first differences (ADF $p < 0.001$; KPSS $p = 0.10$). This raises spurious regression concerns for the level specifications (Models 1–3). Following the two-step procedure of Engle and Granger (1987), we apply cointegration tests to Model 2 residuals, which reject the null of a unit root (ADF $p < 0.001$, test statistic -4.90), supporting a long-run equilibrium interpretation. We note that the KPSS test on the same residuals rejects the null of stationarity ($p = 0.01$), creating an ADF/KPSS conflict; given the strong ADF rejection (well below the 1% critical value) and the known tendency of KPSS to over-reject in the presence of structural breaks, we follow the ADF conclusion but acknowledge this ambiguity (see Appendix A for full results). Pairwise Engle-Granger tests indicate that the spread is cointegrated with BDTI ($p = 0.008$) but not with shadow share ($p = 0.684$) or Brent ($p = 0.140$) individually; it is the joint model specification that achieves cointegration, consistent with the interpretation that freight costs are the dominant long-run driver of the spread level.

Our dynamic Model 4, which includes the lagged dependent variable (coefficient $= 0.964$), provides an implicit error-correction representation that is robust to whether the cointegrating relationship is stable. The near-unit-root autoregressive coefficient implies high persistence and suggests that long-run multipliers from Model 4 should be interpreted with caution; we therefore rely primarily on Models 1–3 for substantive inference. We also estimate all specifications in first differences as a complementary analysis, where the spread change is regressed on event impulse dummies. The first-differenced specifications confirm the key findings: the war impulse is large and significant ($8.1/day, $p < 0.01$), while cap and Biden impulses are economically small.

3 Results and Discussion

Event Study Results. Table 2 presents event study results. The war outbreak generated highly significant abnormal returns: the spread widened by $7.90 on the event day and cumulated to $22.35 over the full window ($t = 7.72$, $p < 0.001$). No subsequent sanction announcement generated statistically significant abnormal returns, suggesting markets had largely anticipated these measures. This is consistent with Kokeny and Kelemen (2025), who find that only surprise events generate significant returns in the Russia-Ukraine context, and with Omar and Lambe (2022), who document that sanctions on oil exporters produce price adjustments whose magnitude depends on surprise content.

A caveat on the estimation window: for post-war events, the estimation window $[-120, -6]$ spans a period of structural change, potentially violating the stable-parameters assumption. Robustness checks with alternative windows $[-3, +5]$ and $[-10, +20]$ yield qualitatively identical conclusions.

Table 2 Event Study Results

Event	Date	Event Day AR	CAR$[-5,+10]$	t-stat	p-value
War Outbreak	2022-02-24	\$7.90	\$22.35***	7.72	0.000
EU Oil Ban	2022-06-03	−\$0.24	−\$2.53	−0.46	0.644
Price Cap	2022-12-05	\$0.76	\$4.24	0.64	0.525
Products Cap	2023-02-05	−\$0.07	−\$0.46	−0.12	0.907
Enforcement	2024-06-12	−\$0.33	−\$4.95	−1.35	0.179
Biden Sanctions	2025-01-10	\$0.33	\$1.29	0.43	0.667

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. CAR computed over $[-5, +10]$ event window. Estimation window: $[-120, -6]$.

Shadow Fleet Evolution. Table 3 documents the evolution of shadow fleet market share. The share rose from approximately 34–36% in 2019–2021 to 62.6% in 2024, then dropped to 45.6% in 2025 following the Biden tanker sanctions.

Table 3 Shadow Fleet Evolution by Year

Year	Total Volume (KT)	Shadow Share (Def B)	Sanct. Share (US)	Unique Vessels
2019	280,430	33.9%	16.2%	857
2020	244,576	35.6%	20.8%	834
2021	263,502	36.2%	22.4%	811
2022	280,411	39.2%	21.5%	794
2023	289,586	52.3%	27.3%	794
2024	284,061	62.6%	29.3%	726
2025	277,942	45.6%	9.5%	760

Notes: Shadow Share (Def B): sanctioned by any jurisdiction OR aged ≥20 years with convenience flag. Sanct. Share (US): OFAC-designated vessels only.

A central consequence of the sanctions regime has been the large-scale reorientation of Russian crude exports from Western to non-Western buyers. Figure 1 documents this shift. Before the war, EU member states collectively received the largest share of Russian seaborne crude. Following the EU embargo, EU-bound volumes collapsed and were replaced by shipments to India and China, consistent with Babina et al (2025) and Levi et al (2023). The reorientation was not costless: the shift toward price-sensitive Asian buyers widened the discount (Kilian et al 2024), and longer voyage distances increased transport costs (Spiro et al 2025). However, it preserved export volumes (Table 1, Panel B). Cardoso et al (2025) formalize this mechanism, showing that the cap incentivized alternative shipping infrastructure rather than binding the price.

The shadow fleet also raises environmental and safety concerns. Shadow fleet vessels are substantially older (mean age 20.4 years versus 12.5 years for non-shadow vessels), with 70% aged 20 years or older compared to 6% of non-shadow vessels. These aging tankers operate outside Western insurance networks, creating uninsured environmental liability in transit areas (see Section 4 for policy implications).

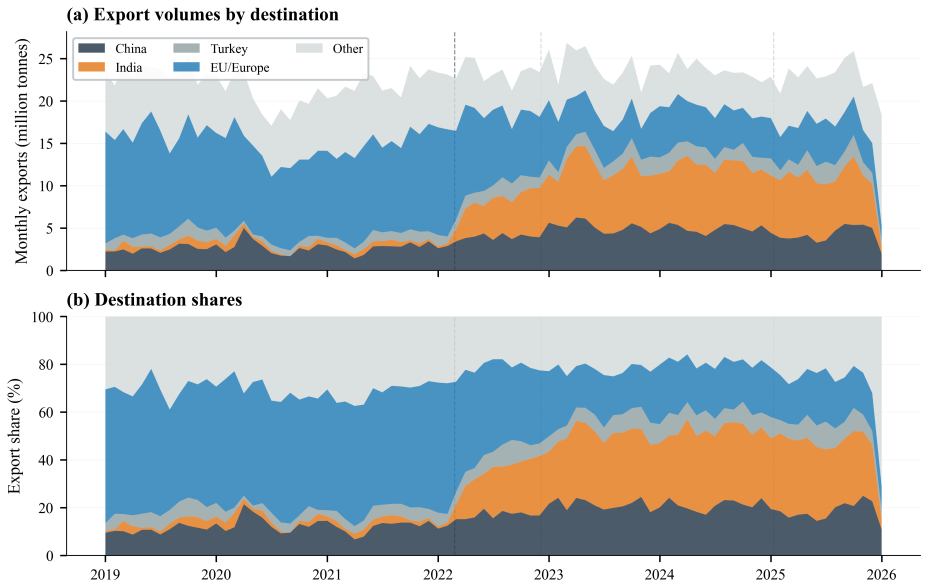


Fig. 1 Trade reorientation of Russian crude oil exports. (a) Monthly export volumes by destination group (million tonnes). (b) Destination shares over time (%). The shift from EU/Europe toward India and China accelerates from mid-2022. Vertical dashed lines mark the war outbreak, price cap, and Biden sanctions. Data cover 25,399 cargo records (January 2019–February 2026).

Regression Results. Table 4 presents our main regression results across four specifications.

The post-war period is associated with a spread increase of \$26.6/bbl across all specifications ($p < 0.01$). The post-cap period shows a spread *decrease* of \$9.22–9.67/bbl. Each percentage point increase in shadow fleet share is associated with a \$0.022 narrower spread ($p < 0.01$). The dynamic Model (4) shows the shadow share coefficient becoming statistically insignificant (+0.001, $p > 0.10$): once the lagged spread absorbs persistent level effects, the short-run *innovation* in shadow share has no detectable association with spread changes, consistent with the shadow fleet's influence operating through slow, structural channels rather than day-to-day responses.

Model (3) reveals that the Post Cap × Shadow interaction is negative (−0.042, $p < 0.01$), indicating shadow fleet expansion was particularly strongly associated with narrower spreads after the cap. The Post Biden × Shadow interaction is positive (0.079, $p < 0.01$), suggesting the relationship reversed following targeted sanctions.

Dollar versus Percentage Spread. When measured as a percentage of Brent, the price cap coefficient becomes statistically insignificant (−0.43, $p > 0.10$), while the war effect remains highly significant (Table 5). This indicates the apparent dollar narrowing reflects declining Brent prices (\$106→\$82) rather than a genuine discount change.

Figure 2 visualizes this mechanism. Panel (a) shows the dollar spread declining in tandem with Brent prices: as Brent fell from approximately \$106/bbl to \$82/bbl, a constant proportional discount would mechanically produce a narrower dollar spread.

Table 4 Determinants of Brent-Urals Spread

Variable	(1) Baseline	(2) Shadow	(3) Interaction	(4) Dynamic
Post War Outbreak	26.62***	26.59***	26.71***	1.25***
	(2.29)	(2.27)	(2.26)	(0.41)
Post Price Cap	−9.67***	−9.22***	−6.76***	−0.48***
	(1.71)	(1.73)	(1.83)	(0.17)
Post Biden Sanctions	−5.89***	−6.11***	−10.61***	−0.21*
	(1.13)	(1.13)	(1.42)	(0.11)
Shadow Fleet Share		−0.022***	−0.015**	0.001
		(0.006)	(0.006)	(0.001)
Post Cap × Shadow			−0.042***	
			(0.014)	
Post Biden × Shadow			0.079***	
			(0.015)	
DXY (std)	−2.23***	−2.17***	−2.20***	−0.19***
	(0.66)	(0.66)	(0.65)	(0.07)
VIX (std)	0.92**	0.95***	0.92***	0.07**
	(0.36)	(0.36)	(0.36)	(0.04)
BDTI (std)	2.67***	2.65***	2.62***	0.10**
	(0.64)	(0.63)	(0.62)	(0.04)
COVID Crash	−2.58**	−2.71**	−2.58**	−0.13
	(1.07)	(1.06)	(1.05)	(0.10)
Spread$_{(t-1)}$				0.96***
				(0.01)
Constant	1.61	2.54**	2.20**	−0.11
	(1.01)	(1.05)	(1.01)	(0.08)
Observations	1,582	1,582	1,582	1,582
R^2	0.851	0.852	0.855	0.993

Notes: Robust standard errors (Newey-West, 10 lags) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. All controls standardized. Shadow Fleet Share measured in percentage points (0–100 scale); coefficient represents the association of a one percentage point increase.

Table 5 Dollar vs. Percentage Spread

Variable	Dollar Spread	Dollar + Brent	% Spread
Post War Outbreak	26.59***	21.31***	18.85***
	(2.27)	(2.09)	(2.25)
Post Price Cap	−9.22***	−5.97***	−0.43
	(1.73)	(1.73)	(1.68)
Post Biden Sanctions	−6.11***	−4.72***	−4.19***
	(1.13)	(1.06)	(1.27)
Shadow Fleet Share	−0.022***	−0.018***	−0.020***
	(0.006)	(0.006)	(0.008)
Brent (std)	—	2.32***	—
		(0.44)	
R^2	0.852	0.866	0.818

Notes: Robust standard errors (Newey-West). *** $p < 0.01$. $N = 1{,}582$. % Spread = (Brent − Urals)/Brent × 100.

Panel (b) confirms this by plotting the spread as a percentage of Brent, which shows no discrete level shift at the cap date. This connects directly to the decomposition framework of Kilian et al (2024) and Spiro et al (2025): the Russian discount reflects buyer bargaining power and transport costs rather than a cap constraint. Wachtmeister et al (2023) reach a complementary conclusion from the supply side, predicting the cap would bind only below $40–50/bbl.

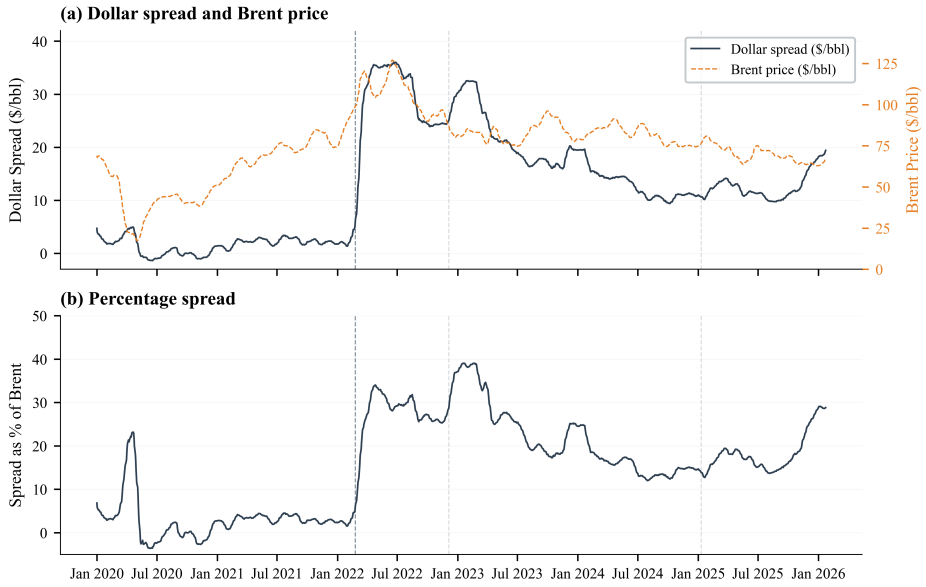


Fig. 2 Dollar spread versus percentage spread: a measurement artifact. (a) The dollar spread (solid) narrows as Brent prices (dashed) decline from $106 to $82/bbl; the regression cap coefficient is −$9.22 ($p < 0.01$). (b) Measured as a percentage of Brent, the cap coefficient becomes statistically insignificant (−0.43pp, $p > 0.10$), indicating the dollar narrowing reflects declining Brent prices rather than a genuine discount change. Vertical dashed lines mark the war outbreak (Feb 2022), price cap (Dec 2022), and Biden sanctions (Jan 2025). Both series are 14-day moving averages.

The spread also narrowed gradually through sub-period decomposition: −$4/bbl early (Dec 2022–Jun 2023), −$12.6/bbl at maturity (Jul 2023–Jun 2024), and −$17.8/bbl late (Jul 2024–Jan 2025), with R^2 improving from 0.852 to 0.919. Shadow fleet share becomes insignificant ($p = 0.95$) as sub-period dummies absorb the gradual mechanism.

War Premium Decay. The war created an initial premium of approximately $30/bbl that decays gradually over time. We estimate four decay functional forms (Table 6): the linear model implies a half-life of 2.8 years, while exponential (2.4 years), log-linear (2.6 years), and quadratic (non-constant) specifications yield consistent estimates. In the linear specification, the shadow fleet coefficient remains significant (−0.021, $p = 0.012$), but in the quadratic and exponential models, shadow fleet share becomes insignificant ($p > 0.20$), suggesting the gradual spread narrowing is largely explained by natural fading of the war shock.

Table 6 War Premium Decay Model Comparison

Model	Half-life (yr)	Shadow Coef	Shadow p	R^2
Linear	2.8	−0.021	0.012	0.783
Quadratic	Non-constant	0.009	0.227	0.843
Exponential (NLS)	2.4	—[b]	—	0.403
Log-linear	2.6	−0.001	0.018	0.787

Notes: All models include post-war observations only (N = 1,022). [b]Shadow share enters as a control in the first-stage OLS; the NLS decay model is estimated on residuals, so no separate shadow coefficient is available.

Figure 3 illustrates this mechanism. Two features stand out. First, the fitted decay captures both the rapid initial narrowing (2022) and the slower convergence during 2023–2024, without requiring any cap-related parameter. Second, the spread diverges upward from the decay path in late 2025, consistent with the Biden tanker sanctions disrupting the smooth decay process. The 2.4–2.8 year half-life implies that roughly half of any war-induced price dislocation dissipates within two to three years through market adaptation, consistent with Hufbauer et al (2007), who document that target economies adjust to restrictions within two to three years, and with Omar and Lambe (2022), who find that oil price effects of sanctions on exporters attenuate over time.



Fig. 3 War premium decay model. The linear decay model (dashed line, half-life = 2.8 years) overlaid on the actual Brent-Urals spread (solid line, 7-day moving average). The dot marks the half-life point. Exponential and log-linear specifications yield half-lives of 2.4 and 2.6 years respectively (Table 6). In the quadratic specification, the shadow fleet coefficient becomes insignificant ($p = 0.227$).

Revenue Impact. To translate spread findings into policy-relevant magnitudes, we estimate revenue losses as Spread × Daily Volume, following the approach of Babina et al (2025). Our central estimate of cumulative post-war Russian revenue losses from

the Brent-Urals discount is $108 billion (Figure 4; $109 billion using daily-frequency aggregation). We construct sensitivity bands by varying the spread assumption ±20% and adjusting for pipeline volumes and FOB pricing. The range spans $87–130 billion, with a FOB-adjusted estimate (subtracting freight costs) of $91 billion and a pipeline-inclusive estimate of $122 billion (Table 7). Losses were highest during the war pre-cap period ($169M/day), then declined to $101M/day post-cap and $71M/day post-Biden, mirroring the war premium decay trajectory.

Table 7 Revenue Loss Sensitivity Analysis ($ Billions)

Period	Base	Low (−20%)	High (+20%)	With Pipeline	FOB-adj.
War Pre-Cap	34.2	27.4	41.0	38.7	30.8
Post-Cap	54.4	43.5	65.3	61.6	45.1
Post-Biden	19.5	15.6	23.4	22.1	14.9
Total (Post-War)	108.1	86.5	129.7	122.4	90.8

Notes: Base: spot spread × seaborne volume. Low/High: ±20% spread variation. Pipeline: includes estimated Druzhba volumes. FOB-adjusted: subtracts estimated freight cost component.

The shadow fleet's contribution to revenue recovery is modest. The $0.51/bbl average counterfactual narrowing translates to approximately $2.2 billion, or 2.2% of total excess losses. This asymmetry between the fleet's rapid expansion (from 40.5% to 63.1% of seaborne volume) and its marginal revenue recovery indicates that the shadow fleet preserved *volume* but could not substantially reduce the *discount*. Wachtmeister et al (2023) predict a similar outcome, showing that circumvention reduces the cap's revenue effect but does not eliminate the structural discount imposed by market segmentation. These estimates are approximations: they assume spot spreads proxy for transaction prices and do not capture side payments, term contract discounts, or barter arrangements. Babina et al (2025), using transaction-level Russian customs data, provide complementary evidence on realized price discounts; our spread-based estimate likely serves as an upper bound since spot spreads may exceed contract-based discounts. The sensitivity analysis above provides bounds for alternative assumptions.

Robustness. We conduct extensive robustness checks; we summarize the key results here, with full details in the Supplementary Materials.

Placebo Event Study. We generate 1,000 random pseudo-event dates in the pre-war period and compute CARs. None produced a CAR exceeding the actual war CAR (empirical $p = 0.000$). The actual CAR lies more than 20 standard deviations beyond the placebo distribution (Figure 5).

First-Differenced Specifications. Re-estimating all models in first differences confirms the key findings (Table 8). The war impulse is large ($8.14/day, $p < 0.01$), while cap and Biden impulses are economically small ($0.54 and $0.59/day). The shadow share change has no contemporaneous association with spread changes. An error correction model (FD-4) confirms negative mean reversion (−0.004, $p < 0.05$) with a half-life of approximately 170 trading days. The low R^2 (0.05) reflects the fact

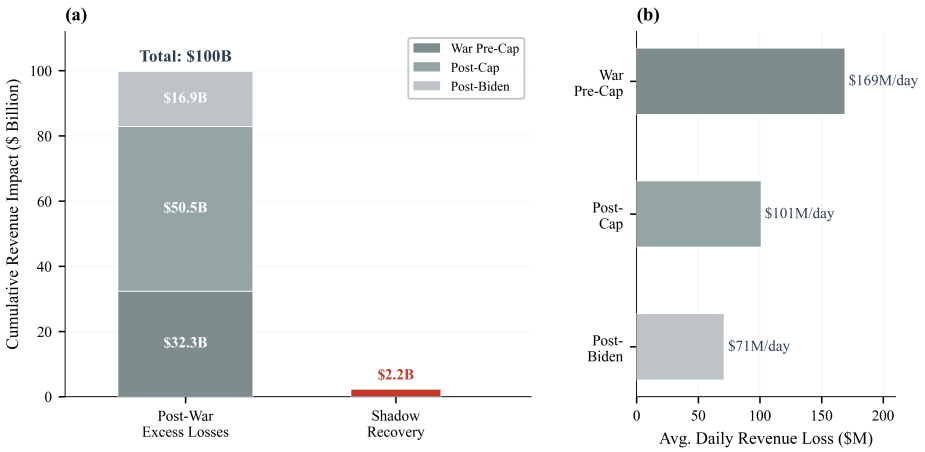


Fig. 4 Revenue loss decomposition. (a) Cumulative post-war excess losses by period ($100B above pre-war baseline; cumulative total including baseline is $109B) versus shadow fleet revenue recovery ($2.2B, 2.2% of excess losses). (b) Average daily revenue loss rate by period, declining from $169M/day (war pre-cap) to $71M/day (post-Biden), consistent with war premium decay. See Table 7 for sensitivity bounds ($87–130B).

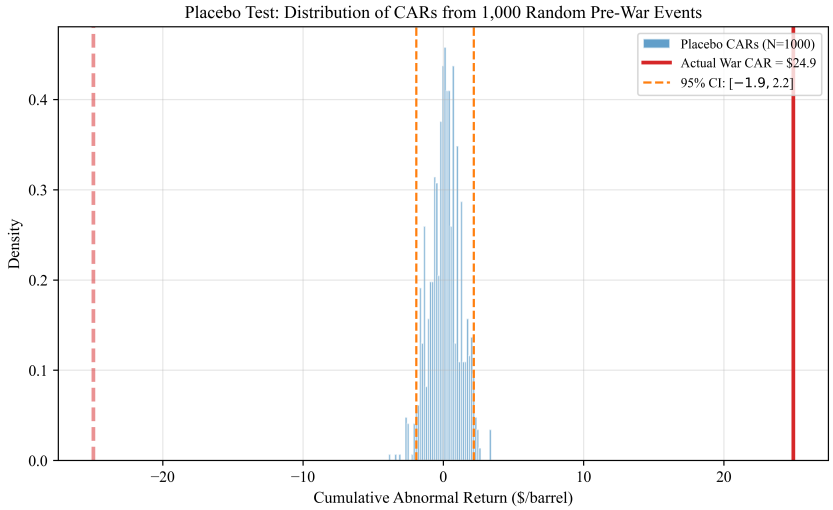


Fig. 5 Placebo Event Study: Distribution of 1,000 Random CARs. *Note:* The figure displays the random-walk benchmark CAR ($24.9); the market-model CAR reported in the text is $22.35.

that daily spread changes are dominated by noise; the level effects that drive the main regression results are, by construction, absent from the first-differenced specification.

Structural Break Tests. Bai-Perron tests identify February 24, 2022 (the war outbreak) as the dominant structural break. Additional breaks cluster around early 2023 and early 2024, consistent with gradual spread narrowing. The PELT-BIC model

Table 8 First-Differenced Specifications

Variable	(FD-1) Baseline	(FD-2) Shadow	(FD-3) Lagged	(FD-4) ECM
War Impulse	8.142***	8.137***	8.178***	8.167***
	(0.163)	(0.164)	(0.171)	(0.169)
Cap Impulse	0.543***	0.542***	0.517***	0.576***
	(0.051)	(0.051)	(0.054)	(0.062)
Biden Impulse	0.594***	0.594***	0.575***	0.553***
	(0.042)	(0.043)	(0.044)	(0.043)
Δ Shadow Share		0.000	0.001	0.001
		(0.001)	(0.001)	(0.001)
Spread_{t-1} [ECM]				−0.004**
				(0.002)
Observations	1,581	1,581	1,581	1,581
R^2	0.051	0.051	0.052	0.054

Notes: Dependent variable: ΔSpread_t. Impulse dummies equal one on the event date only. Controls: ΔDXY, ΔVIX, ΔBDTI (all standardized). FD-3 includes $\text{Shadow Share}_{t-1}$ (level). FD-4 adds Spread_{t-1} as error correction term. *** $p < 0.01$, ** $p < 0.05$.

identifies a break near December 2, 2022, but this coincides with the EU embargo, not solely the cap (Table 9; Figure 6).

Table 9 Bai-Perron Structural Break Tests

N Breaks	Break Dates	Total Cost
1	2022-02-24	67,787
2	2022-03-01; 2023-04-05	17,584
3	2022-03-01; 2023-03-28; 2024-01-23	11,712
6 (PELT-BIC)	2022-03-01; 2022-08-17; 2022-12-02; 2023-03-28; 2024-02-15; 2025-11-03[a]	—

Notes: Dynamic programming algorithm for 1–3 breaks; PELT with BIC penalty for data-driven selection. [a]Identified in the final months of the sample; limited post-break observations warrant caution.

ESPO Spread. Using the Brent-ESPO spread (Pacific exports) as the dependent variable, we find qualitatively similar but smaller coefficients: the war-related increase is \$20.7/bbl (versus \$26.6 for Urals), and the post-cap decrease is \$7.8/bbl. ESPO traded above the \$60 cap in 91.6% of post-cap observations, confirming widespread price cap violations on the Pacific route. The lower R^2 (0.586 versus 0.852) reflects the different dynamics of Asian-bound exports, which are less affected by the EU embargo.

Pacific vs. Atlantic Quasi-DID. A key identification challenge is separating the EU embargo (which applies only to Atlantic/Urals exports) from the price cap (which applies to both routes). We exploit this asymmetry in a quasi-difference-in-differences framework. If the EU embargo drove the post-cap spread narrowing, the Atlantic

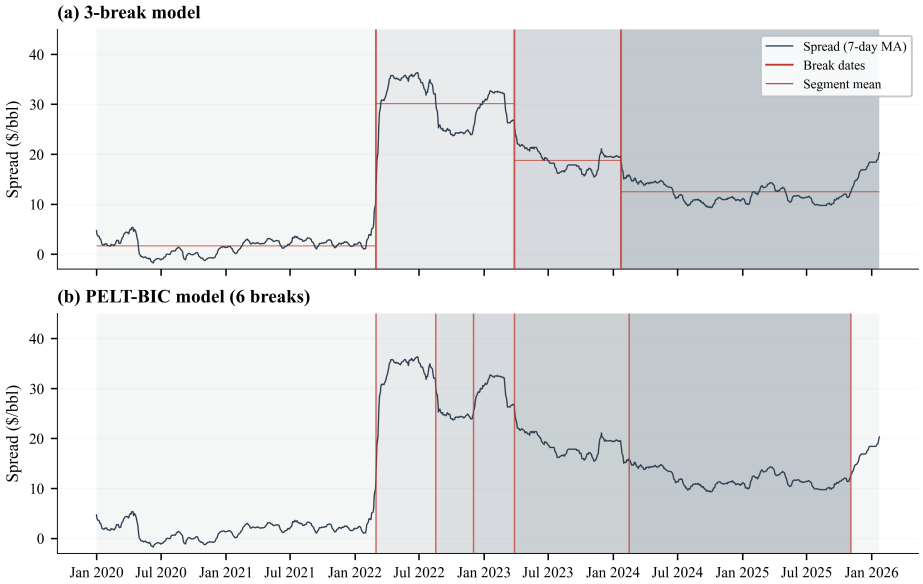


Fig. 6 Structural break regime map. (a) Bai-Perron 3-break model. Horizontal lines indicate segment means; alternating shading distinguishes regimes. (b) PELT-BIC data-driven model identifying six breaks. The December 2, 2022 break coincides with both the EU embargo and the price cap.

(Urals) spread should narrow more than the Pacific (ESPO) spread. We estimate separate regressions for each route and test whether the PostCap coefficient differs. The Atlantic PostCap coefficient is −$7.85 and the Pacific PostCap coefficient is −$6.94; the difference (−$0.91, $p = 0.72$) is not statistically significant (Table 10). We note that this is a failure to reject the null of equal coefficients, not evidence of equality; the large standard error on the difference ($2.54) implies limited power to detect differences smaller than approximately $5/bbl at conventional significance levels. For the PostBiden coefficient, the difference is larger (−$2.96, $p = 0.011$), consistent with targeted tanker sanctions having a stronger Atlantic effect. The Baltic Dirty Tanker Index shows opposite signs: positive for Atlantic (+2.71, $p < 0.01$) but negative for Pacific (−1.52, $p < 0.05$), reflecting the different freight dynamics of the two routes. These results suggest that the post-cap spread narrowing was not driven by the EU embargo alone but rather by a common factor affecting both routes, consistent with the war premium decay interpretation.

Alternative Shadow Fleet Definitions. Using Definition A (sanctioned vessels only), the shadow fleet coefficient becomes insignificant (−0.001, $p = 0.82$). Decomposing into "old shadow" and "new shadow" (post-war entrants), the new shadow coefficient (−0.048, $p < 0.01$) is substantially larger than the old (−0.020, $p < 0.01$), supporting the interpretation that sanctions-induced fleet expansion had a measurable effect. The sensitivity of the shadow fleet coefficient to definition choice (significant under B, insignificant under A) highlights that measurement of the shadow fleet remains an important source of uncertainty in this literature.

Table 10 Pacific vs. Atlantic Quasi-Difference-in-Differences

	Atlantic (Urals)		Pacific (ESPO)		Difference	
Variable	Coef	SE	Coef	SE	Coef	SE
Post Price Cap	−7.85***	(2.06)	−6.94**	(2.70)	−0.91	(2.54)
Post Biden	−6.64***	(1.35)	−3.68***	(0.99)	−2.96**	(1.16)
Shadow Share	−0.029***	(0.009)	−0.015*	(0.008)	−0.014*	(0.008)
BDTI (std)	2.71***	(0.81)	−1.52**	(0.67)	4.23***	(0.83)
R^2	0.639		0.334		0.408	
N	1,022		1,022		1,022	

Notes: Post-war observations only. Controls: DXY (std), VIX (std), BDTI (std). Difference column tests H_0: Atlantic = Pacific. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Additional checks (reported in the Supplementary Materials) include alternative spread measures (Novorossiysk, average), lagged shadow share (1–14 days, stable coefficients), Granger causality tests (no reverse causality), VAR impulse response functions, freight-adjusted spreads, counterfactual simulation, diagnostic tests (RESET $p = 0.92$; max VIF = 8.63), and subsample robustness.

Limitations. Several limitations qualify our findings. First, identification relies on step dummies and event studies rather than quasi-experimental variation; we cannot fully isolate sanctions effects from concurrent macroeconomic shifts. Second, our shadow fleet Definition B includes pre-existing old vessels, not solely sanctions evaders; the sanctioned-only Definition A yields insignificant coefficients, suggesting measurement matters. Third, daily spot prices may not reflect term contract prices or side payments. Fourth, the FOB/CIF mismatch is addressed through BDTI controls but not eliminated. Fifth, non-stationarity in levels means level regression R^2 values may be inflated, though cointegration tests and first-differenced specifications support the key findings. Sixth, our volume data excludes pipeline exports (Druzhba), which continued to supply EU refineries under exemptions.

4 Conclusion and Policy Implications

This study examines G7 oil sanctions on Russia using daily price data and vessel-level cargo records from January 2020 to January 2026. Our central conclusion is that the G7 price cap did not function as a binding constraint on Russian crude oil prices. The war outbreak, not the cap, is associated with the dominant change in the Brent-Urals spread ($26.6/bbl, half-life 2.4–2.8 years depending on functional form). The apparent spread narrowing under the cap is a measurement artifact: insignificant as a percentage of Brent, absorbed by war premium decay, and gradual over two years. A Pacific versus Atlantic quasi-difference-in-differences analysis confirms that the narrowing was not driven by the EU embargo alone. Shadow fleet expansion increased market share from 40.5% to 63.1%, but both the shadow fleet and the spread appear to respond to a common process of market adjustment. Revenue recovered by the shadow fleet ($2.2 billion) is modest relative to total post-war losses ($87–130 billion, central estimate

$109 billion). Targeted vessel sanctions (January 2025) proved more effective, reducing shadow fleet share from 63.1% to 51.7% (period averages).

Six policy implications follow from these findings.

First, **the price cap was ineffective as a binding constraint**, functioning instead as a partial transport embargo (Cardoso et al 2025; Spiro et al 2025). Its primary effect was to incentivize shadow fleet expansion rather than constrain Russian revenues. The EU import embargo and resulting market fragmentation were the more consequential sanctions, consistent with Kilian et al (2024).

Second, **targeted vessel sanctions are more effective** at disrupting evasion infrastructure than price-level adjustments. The Biden designations (January 2025) reduced shadow fleet share by 11.4 percentage points (from 63.1% to 51.7%, period averages), demonstrating that enforcement against specific vessels can disrupt the evasion network. Sustained enforcement, including the October 2025 sanctions on Rosneft and Lukoil and subsequent OFAC designations, appears to be the most effective channel.

Third, **lowering the cap level may not suffice without enforcement reform**. The EU's 18th sanctions package (July 2025) reduced the cap from $60 to $47.60/barrel, with a dynamic mechanism lowering it further to $44.10 in January 2026 (European Commission 2026). Our analysis provides context for evaluating this adjustment: with post-Biden Urals prices averaging $56.55/bbl, a cap at $44.10 approaches the binding range. However, the shadow fleet's persistence (52% share even post-Biden) means that enforcement, not price level, remains the binding constraint. Unless flag-state cooperation improves and insurance attestation compliance is strengthened, the lower cap may simply accelerate further shadow fleet growth, replicating the pattern observed after the original $60 cap.

Fourth, **international coordination on flag-state oversight is essential**. Shadow fleet vessels exploit convenience flags and opaque ownership structures (Rodriguez-Diaz et al 2025; Levi et al 2023). The gap between aggregate fleet data and sanctioned-vessel lists (Definition B versus Definition A) reveals the scale of "gray zone" evasion that falls outside current enforcement mechanisms.

Fifth, **war premiums decay naturally** (half-life $\approx$ 2.4–2.8 years). Policymakers should distinguish sanctions effects from natural market adjustment when evaluating policy success. The cap was implemented when the spread was already narrowing, making its apparent effect indistinguishable from shock dissipation. This suggests that the cap's design, while innovative, addressed a constraint (the price level) that was less binding than the market structure (the embargo-induced segmentation). Future sanctions design should target the structural channels (insurance markets, shipping registries, financial intermediaries) through which evasion operates, rather than adjusting the cap level within a system that can be circumvented.

Sixth, **the aging shadow fleet poses environmental risks** that compound the policy challenge. Shadow fleet vessels average 20.4 years of age (versus 12.5 years for non-shadow), with 70% exceeding the 20-year threshold. These vessels operate without Western P&I insurance, creating uninsured environmental liability in ecologically sensitive transit areas. Policymakers should consider environmental risk as an additional

dimension of sanctions enforcement, potentially using maritime safety conventions and port-state controls to restrict aging, uninsured tankers.

More broadly, our findings illustrate the challenge identified in the sanctions literature: large-economy targets with alternative trading partners can adapt within two to three years (Hufbauer et al 2007; Felbermayr et al 2024). The G7 price cap, designed to balance Russian revenue reduction against global supply preservation, achieved the latter at the expense of the former. The welfare trade-off is non-trivial: a strictly binding cap would have reduced Russian revenues more effectively but risked the supply disruption that policymakers sought to avoid (Wachtmeister et al 2023; Turner and Sappington 2024).

Data Availability

Price data are from Refinitiv Eikon (subscription required; accessed via Refinitiv Workspace). Cargo and vessel-tracking data are from LSEG World Flows (subscription required). Both datasets are subject to commercial licensing restrictions that prevent public redistribution. Sanctioned vessel lists are from FleetLeaks (publicly available at https://fleetleaks.com). VIX data are from the Federal Reserve Bank of St. Louis FRED database (publicly available at https://fred.stlouisfed.org). Replication code (Python scripts for all tables and figures) and processed datasets (with proprietary fields removed) are available from the author upon reasonable request.

Declarations

Competing Interest. The author declares no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Funding. No funding was received for conducting this study.

Author Contribution. Tuğba Ortay Baykal: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization.

Acknowledgments. The author thanks the Refinitiv/LSEG team for data access and the FleetLeaks project for making sanctioned vessel data publicly available.

References

Babina T, Hilgenstock B, Itskhoki O, et al (2025) Russian oil exports under international sanctions. Energy Economics 153:109006. https://doi.org/10.1016/j.eneco.2025.109006

Cardoso DSP, Metulini R, Casarin R (2025) The impact of the price cap on the rerouting of Russian crude oil: A difference-in-differences approach, ca' Foscari University Working Paper

Engle RF, Granger CWJ (1987) Co-integration and error correction: Representation, estimation, and testing. Econometrica 55(2):251–276

European Commission (2026) New dynamic mechanism to lower the price cap on Russian crude oil to $44.10 per barrel. https://finance.ec.europa.eu/news/new-dynamic-mechanism-lower-price-cap-russian-crude-oil-4410-barrel-2026-01-15_en, accessed January 2026

Felbermayr G, Morgan TC, Syropoulos C, et al (2024) Economic sanctions: Evolution, consequences, and challenges. Journal of Economic Literature 62(4):1503–1579

Fernandez-Villaverde J, Li Y, Xu L, et al (2025) Charting the uncharted: The (un)intended consequences of oil sanctions and dark shipping, https://doi.org/10.3386/w33486, nBER Working Paper 33486

Hufbauer GC, Schott JJ, Elliott KA, et al (2007) Economic Sanctions Reconsidered, 3rd edn. Peterson Institute for International Economics, Washington, DC

Itskhoki O, Ribakova E (2024) The economics of sanctions: From theory into practice. Brookings Papers on Economic Activity 55(Fall):425–470

Johnson S, Rachel L, Wolfram C (2023) Design and implementation of the price cap on Russian oil exports. Journal of Comparative Economics 51(4):1244–1252. https://doi.org/10.1016/j.jce.2023.06.003

Kilian L, Rapson D, Schipper B (2024) The impact of the 2022 oil embargo and price cap on Russian oil prices, federal Reserve Bank of Dallas Working Paper No. 2401. Forthcoming, *The Energy Journal*

Kokeny L, Kelemen H (2025) The effects of war sanction announcements on oil prices: An event study. Közgazdasági Szemle 72(1):20–49. https://doi.org/10.18414/KSZ.2025.1.20

Levi I, Katinas P, Myllyvirta L, et al (2023) Shedding light on shadow tankers: Who transports Russian oil 18 months into the invasion? Tech. rep., Centre for Research on Energy and Clean Air

MacKinlay AC (1997) Event studies in economics and finance. Journal of Economic Literature 35(1):13–39

Newey WK, West KD (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55(3):703–708

Omar A, Lambe BJ (2022) Crude oil pricing and statecraft: Surprising lessons from US economic sanctions. International Review of Financial Analysis 83:102314. https://doi.org/10.1016/j.irfa.2022.102314

Rodriguez-Diaz E, Alcaide JI, Endrina N (2025) Shadow fleets: A growing challenge in global maritime commerce. Applied Sciences 15(12):6424. https://doi.org/10.3390/app15126424

de Souza G, Hu N, Li H, et al (2024) (trade) war and peace: How to impose international trade sanctions. Journal of Monetary Economics 146:103572

Spiro D, Wachtmeister H, Gars J (2025) Assessing the impacts of oil sanctions on Russia. Energy Policy 206:114739. https://doi.org/10.1016/j.enpol.2025.114739

Turner DC, Sappington DEM (2024) On the design of price caps as sanctions. International Journal of Industrial Organization 97:103099

Wachtmeister H, Gars J, Spiro D (2023) The price cap on Russian oil: A quantitative analysis, https://doi.org/10.2139/ssrn.4497485, sSRN Working Paper 4497485

Appendix A Unit Root Tests

Table A1 Augmented Dickey-Fuller and KPSS Tests

Variable	ADF Stat	ADF p	KPSS Stat	KPSS p	Conclusion
Levels					
Spread (Baltic)	−2.060	0.261	0.925	<0.01	Non-stationary
Spread (Novo)	−1.711	0.425	1.021	<0.01	Non-stationary
Shadow Share	−3.344	0.013	0.769	<0.01	Inconclusive
Brent	−1.995	0.289	1.099	<0.01	Non-stationary
DXY	−1.650	0.457	0.693	<0.01	Non-stationary
VIX	−4.669	<0.001	0.193	0.019	Inconclusive
BDTI	−2.489	0.118	0.523	<0.01	Non-stationary
First Differences					
Spread Change	−7.093	<0.001	0.120	0.098	Stationary
Shadow Change	−14.336	<0.001	0.036	>0.10	Stationary
Brent Return	−7.552	<0.001	0.061	>0.10	Stationary

Notes: ADF H_0: unit root (reject ⇒ stationary). KPSS H_0: stationary (reject ⇒ non-stationary).

Appendix B Freight-Adjusted Spread

Since Brent is quoted FOB and Urals CIF, the observed spread includes a transport cost component. We estimate the freight component from the pre-war BDTI-spread relationship and subtract it. Results are virtually identical to the baseline because the freight adjustment is a linear function of BDTI, which is already a regression control (Table B2).

Appendix C Variance Inflation Factors

Table B2 Original vs. Freight-Adjusted Spread

Variable	Original	Freight-Adjusted
Post War Outbreak	26.59***	26.59***
Post Price Cap	−9.22***	−9.22***
Post Biden Sanctions	−6.11***	−6.11***
Shadow Fleet Share	−0.022***	−0.022***
BDTI (std)	2.65***	1.54**
R^2	0.852	0.827

Notes: Freight-adjusted spread subtracts estimated freight component from observed spread. $N = 1{,}582$.

Table C3 Variance Inflation Factors (Model 2)

Variable	VIF
Shadow Fleet Share	1.32
Post War Outbreak	8.63
Post Price Cap	4.26
Post Biden Sanctions	1.54
COVID Crash	1.97
DXY (std)	4.07
VIX (std)	2.14
BDTI (std)	2.33

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.tex