NET WORK LISER

Discussion Paper Series

IZA DP No. 18772

July 2026

The Cost of Industrial Destruction: Evidence from Ukrainian Strikes on Russian Refineries

Štěpán Mikula
Masaryk University
and IZA@LISER

Fabio Sabatini
Sapienza University of Rome
and IZA@LISER

The Cost of Industrial Destruction: Evidence from Ukrainian Strikes on Russian Refineries*

Abstract

We study the local economic effects of Ukrainian long-range strikes on Russian oil refineries, combining a manually verified event-level record of refinery strikes with quality-screened daily satellite radiance from NASA Black Marble. The analysis covers 29 large Russian refineries, 22 of which sustain verified direct hits over June 2022–May 2026. Using a monthly staggered difference-in-differences design that compares refineries after their first verified strike with not-yet-struck and never-struck refineries, we find that nighttime radiance falls immediately and persistently after a refinery enters the strike campaign: by roughly 30 percent in the innermost measured ring around the facility and by 15–18 percent within five kilometers, with the effect attenuating until it becomes small at twenty-five kilometers. The innermost ring shows no recovery within the first post-strike year. A complementary daily instrumental-variables design uses directional wind alignment to shift strike incidence and traces the strike-day dynamic at higher frequency. Radiance rises at short horizons, consistent with a fire-related light signature corroborated by independent NASA FIRMS detections, and turns negative at a six-month horizon, with weak-instrument-robust inference. Applying the same satellite product and empirical specification to 106 non-refinery deep-strike targets produces no comparable contraction, weighing against a generic-war-disruption interpretation.

JEL classification

D74, F51, H56, L71, O13

Keywords

conflict economics, industrial destruction, drone warfare, nighttime lights, staggered difference-in-differences, instrumental variables, Russia-Ukraine war

Corresponding author

Fabio Sabatini
fabio.sabatini@uniroma1.it

* We gratefully acknowledge ACLED for providing temporary access to the Partner-Tier data used to construct the strike record. ACLED bears no responsibility for the analysis, interpretation, or conclusions of the paper. Mikula gratefully acknowledges support from the project "Research on Peripheries to Strengthen the Resilience of Czech Society", Reg. No. CZ.02.01.01/00/23_025/0008727, co-funded by the European Union (Ministry of Education, Youth and Sports of the Czech Republic, OP JAK).

1 Introduction

Ukrainian long-range strikes on Russian oil refineries have become a sustained campaign against a key asset of Russia's war economy. Refineries transform crude oil into fuels for domestic use and refined products for export; they operate through large, fixed, geographically identifiable facilities; and refined-product markets are difficult to adjust at short horizons. Hydrocarbon revenues account for a substantial share of federal budget resources, while refining capacity is concentrated in a limited number of industrial sites that cannot be dispersed or replaced quickly (Dabrowski, 2025; Yermakov, 2025; Kaushal and Suess, 2025). Yet the damage caused by these attacks is difficult to assess from official accounts alone: losses to strategic infrastructure are politically sensitive, and reported damage is often downplayed or denied.

This paper estimates the local economic consequences of Ukrainian strikes on Russian oil refineries. We do not quantify the aggregate macroeconomic or fiscal cost of the campaign, and we do not assess its strategic effectiveness. Instead, we ask whether strikes on refineries leave a measurable economic footprint in the local economy around targeted facilities. We first examine the spatial signature of this footprint: whether the decline is strongest near the refinery and attenuates with distance, marking a local contraction of economic activity rather than a broad, region-wide wartime downturn. We then ask whether the disruption persists beyond the immediate fire and emergency-response phase.

The setting raises two empirical challenges. The first is measurement. Timely and verifiable local economic data around Russian strategic infrastructure are not available at the spatial and temporal resolution required for this analysis. Even where aggregate official statistics exist, their interpretation is difficult in an authoritarian wartime economy. Nighttime lights provide an independent alternative: Martínez (2022) shows that autocracies systematically report higher growth than nighttime lights would predict. We therefore measure activity around each refinery using satellite-based nighttime radiance, following the broader literature that uses satellite imagery to study economic outcomes when conventional data are unavailable, unreliable, or too aggregated (Henderson et al., 2012; Donaldson and Storeygard, 2016).

The second challenge is identification. Refinery strikes are not randomly assigned. Target selection and timing plausibly reflect strategic value, operational feasibility, expected damage, Russian defensive capacity, and the evolving reach of Ukrainian long-range drones. A comparison between struck and unstruck refineries, or between periods before and after a strike, therefore, cannot by itself be interpreted as causal. Any empirical design must separate the local economic consequences of being struck from the strategic process that determines which refineries are attacked and when.

We address these challenges with two complementary empirical strategies. The first exploits

the staggered timing of first verified strikes across refineries to estimate the evolution of nighttime radiance after a facility first enters the strike campaign, relative to refineries not yet struck or never struck. This monthly design captures the medium-run consequences of entering the strike campaign, rather than the effect of an isolated attack: subsequent strikes on the same refinery are part of the post-treatment exposure regime. Because first-strike timing may still reflect unobserved features of targeted facilities, we report the full event-study profile and assess sensitivity to violations of parallel trends.

The second strategy operates at the strike-day frequency. It exploits directional wind alignment as an instrument for strike incidence, using day-to-day variation in wind conditions to isolate strike-favorable days. This daily design provides a distinct source of variation with which to trace the immediate and delayed response of radiance around strike days. The dynamic profile is informative about the physical sequence of a strike. If a strike causes fire at the refinery, radiance should initially rise; if it disrupts ordinary productive activity, the initial light signature should later give way to a decline. We validate the fire component of this sequence using independent active-fire detections from NASA FIRMS, which provide a separate satellite record of thermal anomalies around recorded strike dates.

The results point to a substantial and persistent decline in nighttime radiance around refineries that enter the strike campaign. The monthly DiD estimates are organized along two dimensions: time since the first strike and distance from the refinery. The decline begins in the month of the first strike and does not visibly reverse over the horizons we observe. For refineries struck early enough in the campaign to be followed for more than a year, the decline remains large beyond month 13. Pre-strike event-study coefficients are small and jointly insignificant, providing no evidence that refineries about to be struck were already on differential radiance trajectories.

The spatial pattern follows a clear local gradient around the facility. Radiance falls by approximately 30% in the nearest measured ring, between 0.5 and 1 kilometer from the refinery centroid, immediately outside the 500-meter disk excluded because it is most exposed to facility lighting, flaring, and immediate fire signatures. Within 5 km, the aggregate decline is 15–18%. The effect attenuates with distance and is small at 25 km. The ring decomposition places the largest and most persistent losses in the inner industrial zone, while showing that the contraction extends beyond the immediate refinery footprint rather than collapsing into a purely mechanical signature at the point of impact.

The daily IV yields a complementary dynamic profile. On days when wind conditions make strikes more likely, radiance initially rises around the targeted refinery, consistent with fire and emergency response at the strike site. The increase is visible at short horizons and is corroborated by FIRMS active-fire detections around recorded strike dates. At longer horizons, the sign reverses: the same strike-favorable variation is followed by a decline in

radiance around six months after the strike, consistent with a loss of ordinary nighttime illumination after physical damage and partial disruption of refinery operations. Anderson–Rubin confidence intervals support the short-run positive phase and the medium-run negative phase despite moderate first-stage strength. The daily evidence also highlights an important measurement issue: around violent events, the immediate fire signature and the subsequent change in ordinary nighttime illumination are distinct objects, and the quality-screened daily product helps separate them.

The result is not an artifact of a single specification or spatial window. It is visible in the event-study profile, robust to sensitivity analyses for violations of parallel trends, supported by wild-cluster and weak-instrument-robust inference, and follows a clear local gradient around struck refineries. The same satellite product and empirical specification produce no comparable decline around other deep-territory targets. Military targets, fuel depots, other industrial plants, power infrastructure, and transport nodes show no statistically significant contraction in the main 5 km comparison, and the contrast remains visible at the 1 km buffer. The refinery effect is therefore difficult to reconcile with an interpretation based only on proximity to strike activity or generic wartime disruption.

The paper speaks first to the empirical literature on the economic consequences of conflict. Existing work has studied terrorism, bombing, civil war, and military destruction through their effects on output, firms, trade, and long-run recovery (e.g., Abadie and Gardeazabal, 2003; Miguel and Roland, 2011; Korovkin and Makarin, 2023; Riaño and Valencia Caicedo, 2024). We contribute to this literature by studying a contemporary campaign of long-range drone strikes against fixed industrial infrastructure inside an authoritarian economy at war. Refineries are especially useful for this purpose because they are economically important, geographically fixed, and observable from space. The analysis provides systematic evidence on the local economic footprint of drone warfare against industrial infrastructure, including its magnitude, timing, and spatial concentration around targeted facilities.

A smaller set of papers studies economic consequences and behavioral responses linked more directly to Russia's war against Ukraine. This work includes evidence on the disruption of Ukrainian production networks, aggregate activity (Korovkin et al., 2025), and social capital (Korovkin and Makarin, 2023), civilian evacuation during the invasion (Martinez et al., 2026), the effects of the full-scale invasion on global wheat markets (Devadoss and Ridley, 2024), the reorganization of trade, foreign direct investment, and portfolio flows across geopolitical blocs after the invasion (Gopinath et al., 2025), and the short-run macroeconomic consequences of the war for inflation and aggregate economic activity in Central and Eastern Europe (Daianu et al., 2025). These studies focus on the invaded country, international markets, third-country macroeconomic outcomes, or civilian behavior during the war. We instead study the local economic footprint of Ukrainian strikes inside Russia, around the refinery infrastructure targeted by the campaign, providing

systematic evidence on whether the war produces measurable local economic losses within the aggressor's economy itself.

Second, the paper contributes to studies using satellite data to measure economic activity in settings where conventional data are unavailable, unreliable, or too aggregated. Following the nightlights literature (Henderson et al., 2012; Donaldson and Storeygard, 2016), we use nighttime radiance to study changes in satellite-visible activity around strategic infrastructure. The Russian wartime setting makes this approach valuable, but it also illustrates a measurement problem that is likely to arise in conflict applications: violent events generate short-run fire signatures that can contaminate nightlight outcomes at the same spatial and temporal scale at which treatment is defined. By combining a monthly difference-in-differences design with a higher-frequency daily analysis and FIRMS active-fire detections, we distinguish the immediate fire signature of a strike from the subsequent decline in ordinary nighttime illumination.

Third, the paper adapts wind-based identification to a data-poor conflict setting. Closest to our use of weather-driven variation, Mahmood and Jetter (2023) exploit day-to-day wind conditions to study the consequences of U.S. drone strikes in Pakistan. More broadly, the use of wind as a source of quasi-random variation builds on environmental and applied-economics work that exploits wind direction to identify the effects of transported pollutants on mortality, medical costs, crime, and educational outcomes (Deryugina et al., 2019; Anderson, 2020; Herrnstadt et al., 2021; Persico and Venator, 2021). In our setting, wind affects the operational feasibility of long-range drone movement rather than the transport of pollution. This adaptation raises its own exclusion concerns, because weather can affect satellite measurement directly. For this reason, the wind-based estimates are read alongside the other evidence in the paper: the timing of the response, the spatial gradient around struck refineries, the strike-day dynamics, the target-heterogeneity exercises, and independent fire detections from FIRMS.

The remainder of the paper is organized as follows. Section 2 describes the refinery strike campaign and its strategic context. Section 3 presents the data. Section 4 lays out the empirical strategies. Section 5 reports the main results from the monthly and daily designs. Section 6 presents the spatial signature of the response, independent satellite-fire validation, and the convergence of the two designs. Section 7 discusses interpretation and limitations. Section 8 concludes.

2 Background: the refinery strike campaign

Ukrainian strikes on Russian oil refineries are a distinct part of the broader long-range campaign against targets deep inside Russian territory. Unlike attacks on mobile military assets, frontline logistics, or border-zone targets, refinery strikes hit fixed industrial facilities

whose location and surrounding economic activity can be observed consistently over time. The first attributed refinery events in our record occur in mid-2022, but the campaign remains limited through 2023 before expanding in 2024 and accelerating sharply in 2025. In our verified strike record, attributed refinery events rise from a single strike in 2022 and 4 in 2023 to 22 in 2024, 83 in 2025, and 32 in the first five months of 2026. Over June 2022–May 2026, the record contains 142 attributed refinery events, 101 of which are verified direct hits on the facility. By 2025–2026, the campaign spans much of European Russia, including refineries in the North Caucasus, the Moscow region, the Volga refining complex, Bashkortostan, and more distant sites reached as Ukrainian long-range capabilities expand. Figure 1 summarizes this escalation.

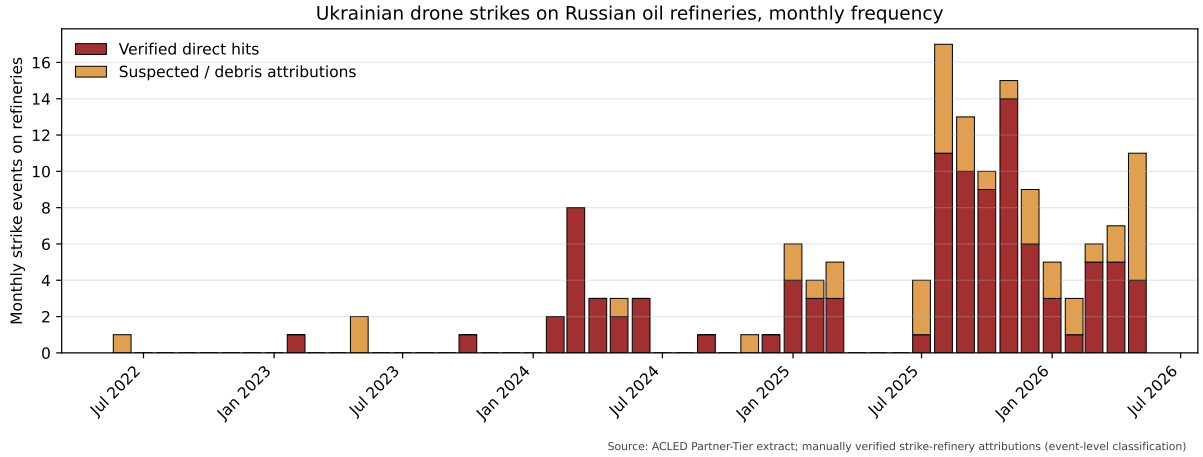


Figure 1: Monthly frequency of Ukrainian drone strikes on Russian refineries

Notes. The figure plots the monthly count of ACLED-recorded strike events on operational Russian oil refineries over the period June 2022–May 2026. Strike events are attributed to refineries through manual, event-level verification of the ACLED record (Section 3.3). The manually verified record contains 142 attributed refinery events over the sample period, of which 101 are verified direct hits; the stacked shading distinguishes verified direct hits from suspected-hit and debris-only attributions. *Source:* authors' calculations from ACLED Partner-Tier extract.

Figure 2 maps ACLED-recorded Ukrainian deep-territory strikes on Russian targets over the sample period, by target type, with each location drawn as a circle whose area reflects the number of strikes it has received. The refineries studied in this paper are one component of a broader deep-strike campaign that also reaches military and airfield targets, other industrial plants, power and energy infrastructure, and ports and shipyards. Within the ACLED air/drone-strike record used here, deep-territory strikes are a small subset of Ukrainian strike activity on Russian territory. Of roughly twenty-seven thousand recorded air/drone-strike events, about 95 percent occur in the three border oblasts of Belgorod, Kursk, and Bryansk and consist primarily of short-range, frontline-adjacent operations against villages and border districts rather than strikes on identifiable strategic infrastructure. We exclude these border-zone events, here and in the analysis, in order to isolate the strategic deep-strike campaign of which refinery strikes form a part.

The Ukrainian refinery campaign is therefore a specific component of a larger and highly

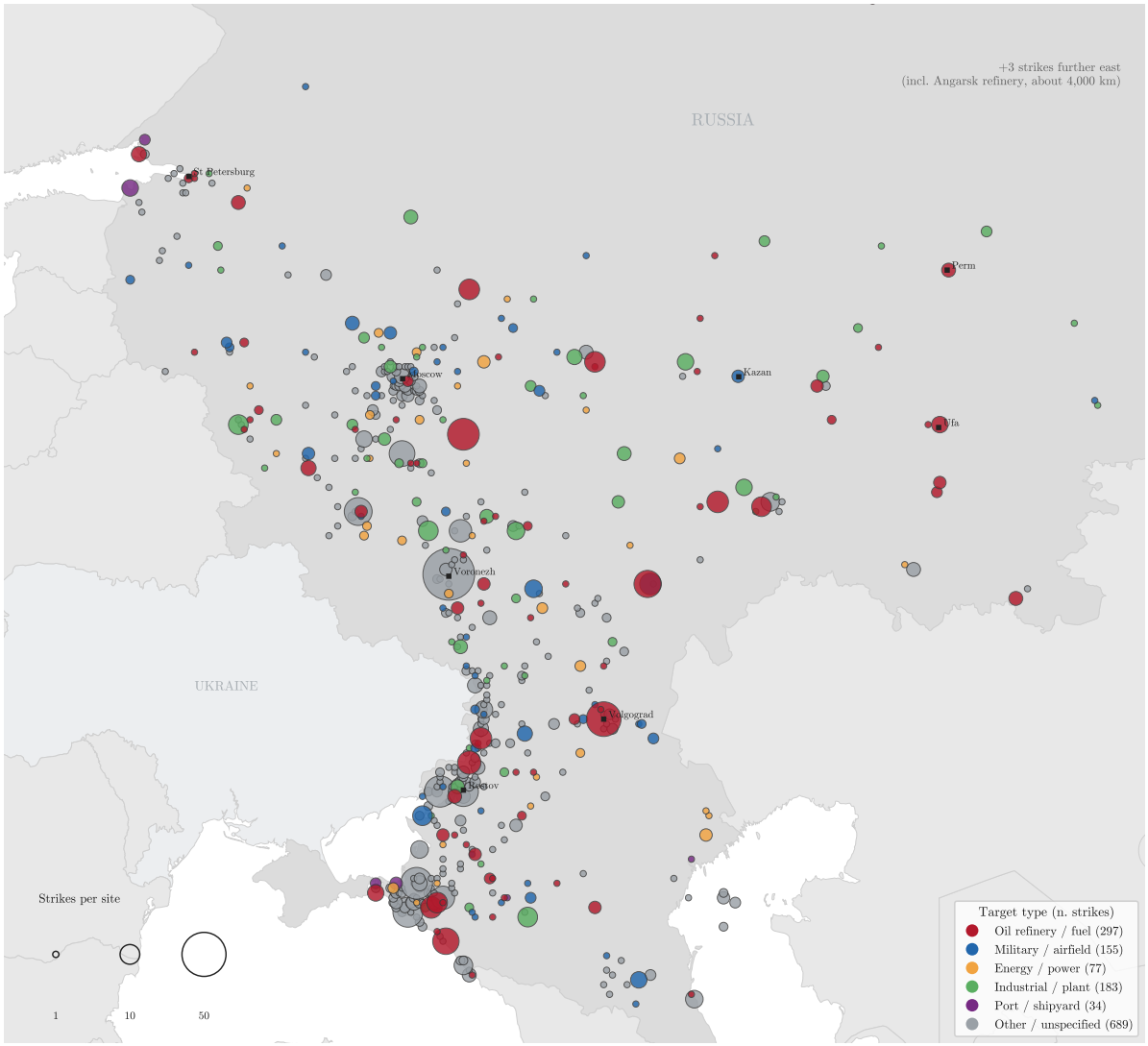


Figure 2: Ukrainian deep-territory drone strikes on Russian targets, by target type

Notes. The map plots the 1,435 Ukrainian air/drone-strike events recorded by ACLED on Russian territory outside the three border oblasts of Belgorod, Kursk, and Bryansk, over February 2022–May 2026. The remaining 25,933 drone-strike events, corresponding to about 95% of the total, occur in those three border oblasts and are excluded: they are short-range, frontline-adjacent tactical operations against villages and border districts rather than strikes on identifiable strategic infrastructure. Strikes are aggregated to distinct locations on a 0.1° grid; the area of each circle is proportional to the number of strikes recorded at that location—the most-struck site carries 69 events—and each location is colored by its predominant target category. Target categories are assigned by keyword search of the ACLED event description; oil refineries, together with closely related fuel infrastructure, are shown in red. Roughly half of deep strikes carry a description specific enough to classify; the remainder are shown as "other / unspecified." The map covers European Russia: a small number of strikes reach further east—including the Angarsk refinery, roughly 4,000 km from the border—and are indicated on the map but not plotted, so as not to stretch the scale. The classification is indicative and serves only to characterize the geographic setting; it is not used in the empirical analysis. *Source:* authors' calculations from ACLED Partner-Tier extract; country borders from Natural Earth.

asymmetric exchange of fire. Appendix Figure 7 maps strike events in both directions: Russian strikes on Ukraine, and Ukrainian strikes on Russia and on Russian-occupied Ukrainian territory. Russian strikes cover much of Ukrainian territory, while Ukrainian strikes concentrate along the shared border and extend, more sparsely, into European

Russia. The number of recorded Russian strikes is roughly three times the number of recorded Ukrainian strikes. This ratio understates the true asymmetry in the scale of fire because ACLED codes discrete events rather than the quantity of ordnance delivered: a sustained Russian artillery bombardment and a single Ukrainian drone sortie each enter as one event. We use the comparison only to situate our object of study. Against this broader backdrop, refinery strikes form a strategically coherent subset of Ukrainian deep-territory operations.

The economic logic of targeting refineries follows from their position in the Russian war economy. Oil and gas revenues remain a central component of federal budget resources, despite the decline in their fiscal weight relative to the pre-war period and the pressure created by sanctions, price caps, and lower export revenues (Dabrowski, 2025; Yermakov, 2025; KSE Institute, 2025). Refineries are also central to domestic fuel supply and to the production of exportable refined products. Unlike upstream extraction, refining capacity is concentrated in a limited number of large industrial facilities with fixed locations, specialized units, and long repair chains. Damage to a crude distillation unit, hydrocracker, vacuum unit, storage tank, or loading facility can reduce effective capacity even when the plant is not destroyed outright. This makes refineries a natural target in a campaign aimed at raising the economic cost of the war and disrupting Russia's capacity to sustain military operations (Kaushal and Suess, 2025; Stockholm Institute of Transition Economics, 2025).

The campaign reflects the rapid expansion of Ukraine's long-range one-way attack-drone capabilities. Public and policy reporting describes the growing use of domestically produced or adapted fixed-wing platforms in attacks against Russian industrial and military targets (Kaushal and Suess, 2025). These systems are reported to be less costly and more scalable than cruise missiles, while carrying modest payloads over long distances. The logic of the refinery campaign is therefore unlikely to be the destruction of an entire facility in a single strike. It is more plausibly a strategy of damaging vulnerable units, interrupting operations, raising repair needs, and imposing defensive burdens on Russia. Repeated attacks on the same refinery are common in our data, consistent with a campaign in which facilities enter a regime of recurrent exposure, disruption, and repair rather than experiencing a single isolated shock.

Weather is operationally relevant for long-range drone missions. One-way fixed-wing drones often travel hundreds, and in some cases more than a thousand, kilometers before reaching fixed targets. Over such missions, wind can affect range, ground speed, flight time, fuel consumption, trajectory control, and exposure to air defense. Engineering work on fixed-wing UAV guidance and path planning shows that wind conditions affect ground speed, path following, and flight time, providing the operational basis for treating directional wind alignment as a feasibility shifter (Voskuijl et al., 2020; Stastny and Siegwart, 2019; Luo et al., 2018). A tailwind along the direction of travel can increase effective range or

reduce flight time; headwinds and strong crosswinds can increase energy requirements or degrade flight stability.

This operational dependence on wind motivates the daily instrumental-variables design developed in Section 4. The empirical idea is not that we observe launch sites or flight paths. Rather, day-to-day variation in winds favoring movement from Ukraine toward a target shifts strike feasibility, both by affecting whether a mission is attempted and by changing the probability that a launched drone reaches the refinery. We therefore use directional wind alignment as a source of strike-day variation, while treating richer weather conditions such as precipitation, snow, and freezing temperatures as diagnostic controls rather than as part of the preferred instrument.

3 Data

Reliable local measurement is a central constraint in this setting. Russian official industrial series provide aggregate information, but they do not offer timely, independently verifiable measures of activity at the spatial scale required here: the industrial perimeter surrounding individual refineries and its evolution around verified strike dates. The empirical analysis, therefore, combines four data streams collected independently of Russian statistical authorities: nighttime radiance from the Visible Infrared Imaging Radiometer Suite (VIIRS), a curated record of drone strikes from the Armed Conflict Location and Event Data project (ACLED), daily meteorological reanalysis from ERA5-Land, and active-fire detections from NASA's Fire Information for Resource Management System (FIRMS).

3.1 Refinery panel

The baseline refinery panel contains 29 major operational Russian oil refineries located within the internationally recognized borders of the Russian Federation as of January 2022. Twenty-two refineries sustain at least one manually verified direct hit during the sample period, while seven are never struck. We refer to the former as treated refineries and to the latter as never-treated refineries. A broader panel variant, used to probe sensitivity to the attribution criterion, additionally treats refineries whose attribution rests on suspected hits or falling debris; it contains 24 treated and 8 never-struck units. Section 3.3 describes the verification procedure, and Appendix B reports the full accounting from the 33-refinery master list to the two analysis panels. In the staggered DiD design, the comparison group is not limited to the never-treated units: depending on the cohort and year-month, it also includes refineries that have not yet received their first verified strike.

The panel is compiled from open and public sources. Refinery identities and locations are drawn primarily from Wikidata and Wikipedia, supplemented by the Global Energy Monitor oil-infrastructure tracker and related open facility databases. Coordinates are

cross-checked against OpenStreetMap industrial-facility tags and satellite imagery when needed. Treatment risk is geographically and strategically heterogeneous across the panel. Most never-treated refineries lie deep in Western and Eastern Siberia and the Russian Far East, far from the front and from any struck facility; one exception is Ufimsky, in the Ufa refining cluster, which sits next to treated refineries yet remained unstruck during the sample period and enters the broad panel only.

Appendix A reports the refinery list, coordinates, treatment status, first-strike date, and event counts. The geographic distribution of refineries across the baseline and broad panels is shown in Appendix Figure 6; the headline estimate is robust to restricting the control group to western refineries only (Appendix G.4).

3.2 Nighttime radiance

We measure local economic activity using satellite nighttime radiance recorded by the Visible Infrared Imaging Radiometer Suite (VIIRS), whose Day/Night Band detects low-light emissions at night. For both identification strategies, the outcome is nighttime radiance measured with NASA's daily Black Marble product, which applies a lunar bidirectional-reflectance-distribution-function correction and pixel-level quality screening at the daily frequency (Román et al., 2018). We retain only pixels whose mandatory quality flag indicates a high-quality retrieval. This screen removes or flags observations affected by clouds, snow, stray light, and other retrieval problems at the pixel-day level. For the monthly difference-in-differences design, we aggregate the quality-screened daily series to the refinery-month; for the daily instrumental-variables design, we use the daily series directly. Daily screening matters here for a conflict-specific reason: strike-induced fires can raise measured radiance in the days around an attack, while our object of interest is the subsequent change in ordinary nighttime illumination.

We construct spatial buffers around each refinery centroid at radii of 1, 2, 5, 10, and 25 km. Every buffer excludes the innermost $(0, 0.5]$ km disk around the centroid, which is dominated by the facility's own lighting and gas flaring and is most exposed to immediate fire signatures. The cumulative buffer at distance d therefore averages radiance over all pixels located more than 0.5 and up to d kilometers from the centroid; the 1 km buffer coincides with the $(0.5, 1]$ km ring. Non-overlapping rings $(d_{k-1}, d_k]$, constructed between adjacent buffer radii, are obtained from pixel-count-weighted differences of the cumulative buffers. The cumulative buffers are the headline measure; the ring decomposition is reported as a spatial diagnostic to distinguish a tight industrial signature from a broader urban-economic spillover.

3.3 Strike data

The treatment record is constructed from ACLED, accessed under the Partner-Tier license.[1] We restrict the search to events occurring within the internationally recognized territory of the Russian Federation as of January 2022. This excludes all Ukrainian territory outside those borders, including Crimea and Sevastopol, which Russia claimed to annex in 2014, and territories in Donetsk, Luhansk, Zaporizhzhia, and Kherson oblasts, which Russia claimed to annex in 2022. All of these territories remain internationally recognized as part of Ukraine.

We retain ACLED events whose sub-event type is *Air/drone strike* or *Shelling/artillery/missile attack* and whose textual description (the ACLED notes field) identifies an oil refinery, refinery complex, or refinery storage facility as the target. We include the second sub-event category because long-range attacks involving drones, missiles, or mixed strike packages can appear under broader shelling/artillery/missile labels. Each candidate event is attributed to a refinery in our panel by name-based matching of the description, followed by manual, event-by-event verification of the full description. We deliberately avoid proximity-based matching: several refineries lie within a few kilometers of air bases and other frequently struck installations, so a radius rule can assign nearby military strikes to the refinery and misdate treatment onset, as in the Saratov–Engels-2 and Ryazan–Dyagilevo cases. From the verified description, we classify each attributed event as a *verified direct hit*, when the description reports an explosion, fire, or damage at the facility itself; a *suspected hit*, when the facility is described as a possible or intended target without confirmed damage; or *falling debris*, when debris from an intercepted drone falls on or near the facility.

The resulting treatment record contains 142 attributed refinery events: 101 verified direct hits, 27 suspected hits, and 14 debris events, spanning June 2022 through May 2026. Most treated refineries are struck repeatedly. Treatment in the staggered DiD design is defined by the month of the first verified direct hit; the broad panel variant dates treatment at the first attributed event of any class. Subsequent strikes on the same refinery do not redefine the treatment cohort. They are part of the post-treatment exposure in the monthly design and motivate the daily IV analysis. Appendix B documents the verification protocol and the full accounting from raw candidate events to the analysis panels.

3.4 Wind instrument

The instrument is a daily measure of whether wind direction favors movement from Ukraine toward each refinery. Because actual drone launch sites and flight paths are unobserved,

[1] ACLED Partner-Tier data cannot be redistributed at the raw event level. Our analysis uses derivative indicators constructed from the licensed event records and aggregated to the refinery-month or refinery-day level.

the instrument is not intended to reconstruct operational routes. For each refinery, we define a reproducible approach bearing from Ukraine toward the target and measure the angular alignment between that bearing and the daily wind direction.

The baseline bearing θ_i is the great-circle bearing from the nearest point on the January 2022 internationally recognized Ukraine–Russia border to refinery i. It is a fixed geographic convention and is not intended to identify a launch site or a flight path. We anchor it on the January 2022 border convention rather than on the moving front line or the boundary of occupied territory, which shift endogenously with the war and are not consistently observed. This choice does not materially affect the results: the instrument depends on the *direction* from Ukraine to each refinery rather than on distance, and a bearing anchored on late-2022 Ukrainian-controlled territory yields a comparable first stage and the same reduced-form pattern—a short-run positive (fire) response followed by a negative response at six months (Appendix D).

Daily horizontal wind components are taken from the ERA5-Land reanalysis, which provides gridded meteorological fields at approximately 9 km spatial resolution (Muñoz-Sabater et al., 2021). We average wind vectors over multiple grid cells along the broad direction from Ukraine to the refinery to reduce dependence on any single reanalysis grid cell and to capture large-scale wind conditions relevant for a long-distance flight. This averaging is a measurement choice, not an attempt to recover the operational path of any specific drone.

The baseline instrument is the cosine of the angular deviation between the approach bearing θ_i and the daily wind direction, $Z_{it} \equiv \cos(\phi_{it} - \theta_i)$, where ϕ_{it} denotes the bearing toward which the daily wind vector points. Positive values indicate winds aligned with movement from Ukraine toward refinery i; negative values indicate winds opposed to that direction. The variable is dimensionless and bounded in $[-1, 1]$, equal to $+1$ when the daily wind points exactly along the approach bearing and -1 when it points exactly against it. We use it directly, without standardization, so that the first-stage coefficient measures the change in daily strike probability as alignment moves from orthogonal to fully favorable.

We discuss the exclusion-restriction logic for this instrument in Section 4. Here we emphasize three features of the construction. First, all inputs are observable from public meteorological data. Second, the baseline instrument captures direction rather than wind speed, reducing the influence of broader storm systems that may affect satellite measurement directly. Third, richer weather conditions such as precipitation, snow, and freezing temperatures are used as diagnostic controls rather than as part of the preferred instrument, because they raise more direct concerns about measurement of nighttime radiance. Appendix I reports a robustness check based on a magnitude-weighted tailwind component, wind speed$_{it} \times Z_{it}$, which combines directional alignment with wind intensity.

3.5 Fire validation

To validate strike timing with an independent satellite source, we use active-fire detections from NASA FIRMS, which provides global active-fire records from the VIIRS and MODIS instruments at 375 m and 1 km resolution, respectively, with a typical revisit cycle of multiple times per day (Schroeder et al., 2014). We restrict to detections within 5 km of each treated refinery and within ±3 days of each ACLED-recorded strike date.

FIRMS serves two roles in the analysis. First, it provides an independent check on the timing of thermal activity around recorded strikes. We do not interpret the frequency of co-located detections as the share of strikes that produce fires: FIRMS records discrete active-fire detections against the background of continuous refinery flaring, whereas nighttime radiance is a continuous measure of light emissions. The validation exercise, therefore, relies on timing rather than levels: active-fire anomalies should rise around recorded strike dates if the strike record captures real thermal events. Section 6.2 shows that these anomalies spike sharply at the strike date relative to each refinery's own baseline. Second, FIRMS provides a physical benchmark for interpreting the strike-day movement in nighttime radiance. In Section 6, we use active-fire timing to distinguish the immediate fire signature of a strike from the subsequent decline in ordinary nighttime illumination.

4 Empirical strategy

Ukrainian strikes on Russian refineries are not randomly assigned. Facilities are targeted because they are economically and strategically relevant, because Ukrainian long-range capabilities expand over time, and because the feasibility of reaching a given target varies with operational conditions, Russian air defense, and weather. A simple comparison between struck and unstruck refineries, or between a refinery before and after it is hit, therefore, cannot by itself be interpreted as causal. It would combine the consequences of being struck with the factors that made a facility more likely to enter the campaign. The empirical design starts from this constraint. A monthly design exploits the staggered timing of first strikes among refineries exposed to the campaign, comparing treated refineries to not-yet-struck and never-struck refineries observed in the same year-month. A daily design exploits higher-frequency variation in strike incidence generated by wind conditions.

The two sources of variation serve different empirical purposes. The monthly difference-in-differences design estimates the trajectory of nighttime radiance after a refinery first enters the strike campaign. This design does not require first-strike timing to be randomly assigned, but it does require a parallel-trends assumption: in the absence of treatment, earlier-treated refineries would have followed local radiance trajectories parallel to those of later-treated and never-treated refineries, conditional on refinery and year-month fixed effects. We assess this assumption through the event-study profile and through sensitivity

checks for violations of parallel trends. The daily instrumental-variables design exploits day-to-day changes in the alignment of wind with the direction from Ukraine toward each refinery. These changes shift the feasibility of a strike occurring at a given refinery on a given day and provide variation in strike incidence that is distinct from the strategic selection of targets. Because weather can also affect satellite measurement, the exclusion restriction is not taken for granted; below we discuss the relevant measurement channel and use horizon-specific estimates, weak-instrument-robust inference, and independent fire detections to assess the interpretation.

This distinction matters for interpretation. The monthly DiD captures the cumulative local effect of entry into the strike campaign. Subsequent strikes on the same refinery are part of the post-treatment exposure rather than separate treatment events in that design. The daily IV instead traces the response of local radiance to strike-day variation induced by wind, allowing us to separate the immediate brightening associated with active combustion from the subsequent decline in ordinary nighttime illumination. NASA FIRMS active-fire detections provide an independent satellite-based validation of the physical fire component of this sequence. Section 5 reports the two sets of results in turn.

4.1 Monthly DiD: campaign exposure

The monthly design estimates the local radiance trajectory that follows a refinery's first verified strike. Treatment is defined by entry into the refinery-strike campaign: once refinery i receives its first verified strike, Post_{im} equals one in that month and in all subsequent months, so that the estimated parameter is the cumulative effect of campaign exposure rather than the marginal effect of an individual strike.

The primary spatial scale is the cumulative 5 km buffer around each refinery. This scale captures the refinery and its immediately surrounding industrial perimeter, while remaining local enough to distinguish facility-level disruption from broader regional shocks. We also report the full spatial gradient across smaller and larger buffers, together with non-overlapping rings, so that the headline estimate is not evaluated in isolation. Temporally, the headline aggregate uses all months from the first verified strike onward. Equation (1) reports the month-by-month event-study profile, with endpoint bins at $k \leq -6$ and $k \geq +12$. For the spatial-dynamic tables, we also report a coarser descriptive decomposition of the post-treatment period into $[0, 3]$, $[4, 6]$, $[7, 12]$, and $[13, +)$ months.

Let m denote calendar year-month and let k denote event time in months relative to refinery i's first verified strike. We estimate the event-study specification

$$\ln Y_{im} = \alpha_i + \lambda_m + \sum_{k \neq -1} \beta_k D_{im}^k + \varepsilon_{im}, \tag{1}$$

where Y_{im} is mean radiance in the relevant buffer around refinery i in year-month m,[2] α_i are refinery fixed effects, λ_m are year-month fixed effects, and D_{im}^k are event-time indicators. The omitted category is $k = -1$, the month immediately preceding the first verified strike. Event time is binned at the endpoints, with $k \leq -6$ and $k \geq +12$. For never-struck refineries, all event-time indicators are equal to zero.

The year-month fixed effects compare refineries within the same year-month. They absorb all shocks common to refineries in a given month, including aggregate war dynamics, sanctions, oil-market conditions, common seasonality, and changes in satellite observation conditions. The remaining identifying variation comes from differences in strike timing across refineries within the same year-month, comparing treated refineries to refineries that are not yet struck or are never struck during the sample period.

For the aggregate campaign-exposure estimate, we replace the event-time indicators with a single post-treatment indicator:

$$\ln Y_{im} = \alpha_i + \lambda_m + \beta_{\text{post}} \text{Post}_{im} + \varepsilon_{im}. \tag{2}$$

We report $\exp\{\beta_{\text{post}}\} - 1$ as the implied percentage change in local radiance.

Because treatment timing is staggered, the aggregate TWFE coefficient should be read as a summary of the post-entry radiance path under the maintained parallel-trends assumption, rather than as a separately identified cohort-specific average treatment effect. Treatment-effect heterogeneity across cohorts and event times can affect TWFE weights in staggered designs. We therefore report the full event-study profile, use never-struck and not-yet-struck refineries as comparison units, and assess whether the aggregate is driven by particular facilities or timing cohorts through leave-one-refinery-out and leave-one-cohort-out exercises.

Identification relies on parallel trends: conditional on refinery and year-month fixed effects, the timing of a refinery's first strike must be uncorrelated with unobserved shocks to its local radiance. The event-study coefficients provide a direct diagnostic, showing whether refineries about to be struck were already on different radiance trajectories before treatment.

Inference is conducted at the refinery level. Standard errors are clustered by refinery throughout. Because the number of clusters is modest ($G = 29$ refineries), we complement asymptotic inference for the headline aggregate with wild cluster bootstrap-t p-values, imposing the null and using both Rademacher and Webb weights. We assess sensitivity to violations of parallel trends using the relative-magnitudes framework of Rambachan and

[2] In log specifications, monthly mean radiance is floored at 0.01 before taking logs. The aggregate 5 km outcome is rarely affected by this floor; Appendix G.3 reports Poisson Pseudo-Maximum-Likelihood estimates in radiance levels, which require no such floor.

Roth (2023).

4.2 Daily IV: directional wind alignment

The daily IV exploits the operational sensitivity of long-range drone missions to wind direction. Long-range propeller-driven UAVs must cover hundreds, and in some cases more than a thousand, kilometers before reaching fixed targets. Over these distances, winds aligned with the direction of travel raise the probability of a successful strike on a given target—shortening flight time and keeping it within reach—while opposed winds lower it. The strike-days entering the first stage are overwhelmingly drone-related, so the operational channel is aligned with the treatment variation used in the daily design.[3] We do not observe drone launch sites or realized flight paths, and we do not attempt to reconstruct them. Instead, we construct a reproducible bearing from Ukraine toward each refinery and use daily variation in wind alignment with that bearing as a source of variation in strike incidence.

For each refinery i, the approach bearing θ_i is the great-circle bearing from the nearest point on the January 2022 internationally recognized Ukraine–Russia border to the refinery, as described in Section 3.4; Appendix C reports the approach bearing for each refinery. This choice defines a stable and reproducible direction from Ukraine toward each target. It is not an assumption about the exact launch site, the realized flight path, or the location of the military front on a given day. A time-varying frontline-based bearing would require imposing a daily map of territorial control and would introduce additional discretion into the construction of the instrument. The January 2022 border, therefore, provides a transparent baseline for measuring directional wind alignment. The results are not sensitive to this convention: anchoring the bearing on late-2022 Ukrainian-controlled territory instead changes the directional-alignment measure very little and leaves the first-stage and reduced-form patterns essentially unchanged (Appendix D).

For each day t, we use ERA5-Land 10 m wind fields to measure the direction toward which the wind blows along the broad direction from Ukraine to the refinery. The baseline instrument is the cosine of the angular deviation between the daily wind direction and the approach bearing:

$$Z_{it} \equiv \cos(\phi_{it} - \theta_i). \tag{3}$$

Here ϕ_{it} denotes the bearing toward which the daily wind vector points. The variable Z_{it} is bounded in $[-1, 1]$. Higher values indicate winds more closely aligned with movement from Ukraine toward refinery i, while lower values indicate winds more opposed to that

[3] In the baseline panel, 93 of 94 strike-days entering the first stage are drone/UAV attacks; in the broader panel, 128 of 129 are drone or mixed drone-package attacks. Excluding the single missile-only strike-day leaves the first-stage F-statistic essentially unchanged: 11.96 to 11.97 in the baseline panel and 10.14 to 10.17 in the broader panel. The reduced-form fire-then-damage pattern is unchanged.

direction. The instrument varies across refineries because θ_i differs across targets, and it varies over time because wind direction changes from day to day. In robustness checks, we also consider a magnitude-weighted tailwind component that multiplies Z_{it} by wind speed.

The outcome at horizon h is a centered seven-day mean of log daily radiance around day $t + h$:

$$\bar{Y}_{i,t,h} \equiv \frac{1}{7} \sum_{\tau=-3}^{3} \ln(\text{rad}_{i,t+h+\tau}) \,. \tag{4}$$

This construction reduces the high-frequency noise of daily Black Marble observations, including cloud-related missingness, overpass timing, lunar-correction residuals, and other single-day retrieval noise. The window is fixed at seven days, applied symmetrically around each horizon, and used uniformly across all reduced-form and 2SLS estimates. Thus, an estimate reported at $h = 180$ should be read as the response in a seven-day window centered 180 days after the strike day.

The estimation sample is the post-first-strike panel of treated refineries. This is the regime in which repeated strikes occur and in which daily wind conditions can shift the timing of additional attacks on a refinery already drawn into the campaign. The first stage relates the binary strike-day indicator to directional wind alignment:

$$\text{Strike}_{it} = \alpha_i + \mu_{m(t)} + \rho_{y(t)} + \omega_{d(t)} + \pi \, Z_{it} + \eta_{it}, \tag{5}$$

where α_i are refinery fixed effects, $\mu_{m(t)}$ are month-of-year fixed effects, $\rho_{y(t)}$ are year fixed effects, and $\omega_{d(t)}$ are day-of-week fixed effects. The reduced form replaces Strike_{it} with the horizon-specific outcome $\bar{Y}_{i,t,h}$. The 2SLS specification instruments Strike_{it} with Z_{it}. Standard errors are clustered at the refinery level.

In the baseline panel, the first-stage F-statistic on directional wind alignment is 12.0; in the broader panel variant it is 10.1. Because the effective first stage varies across horizons with outcome availability, we report Anderson–Rubin confidence intervals throughout the IV analysis (Anderson and Rubin, 1949). A directional-symmetry check supports the operational interpretation: a binary headwind indicator, defined by winds opposed to the approach bearing, predicts strike incidence with a negative coefficient of similar magnitude to the positive directional-alignment effect. If directional alignment affects mission feasibility, winds opposed to the target direction should reduce strike incidence.

The main exclusion concern is that wind may affect measured radiance directly through satellite observation conditions. Wind systems can move aerosols, clouds, and water vapor, and some wind directions may be associated with clearer or noisier scenes independently of any strike. The horizon structure of the IV helps discipline this concern. A contemporaneous measurement channel is most relevant for short horizons, especially around the strike day, where we therefore rely on the sign, timing, and independent FIRMS validation

of the fire signal. At longer horizons, such as $h = 180$, the outcome is measured months after the wind realization that shifted strike incidence and is not subject to the same contemporaneous satellite-measurement channel. The fire-then-damage pattern is also informative: a fixed direct effect of wind on measurement would not naturally generate an immediate positive response followed by a delayed negative response at the same spatial scale.

We retain directional wind alignment as the preferred daily instrument after evaluating alternative weather-based specifications that add wind speed, crosswind, heavy precipitation, snowfall, freezing conditions, and target-level weather controls. These richer specifications do not strengthen the first stage, and using them as the preferred instrument would raise additional exclusion concerns because several weather variables can directly affect satellite retrieval or surface reflectance. Appendix I reports the underlying first-stage audit. It also reports a separate robustness check based on a magnitude-weighted tailwind component, which produces the same fire-then-damage profile. We therefore keep the baseline instrument focused on directional alignment, the component of wind most closely tied to the operational feasibility channel and least directly tied to broader weather conditions that may affect nighttime-radiance measurement.

5 Main results

This section reports the empirical results in three steps. Section 5.1 presents the monthly difference-in-differences estimates, focusing on the event-study profile, the aggregate post-entry effect, the spatial gradient, and sensitivity diagnostics. Section 5.2 turns to the daily instrumental-variables design, which uses directional wind alignment to study strike-day dynamics at a higher frequency. Section 5.3 then compares refinery strikes with other categories of deep-territory targets.

5.1 Cumulative campaign effect: monthly DiD

The monthly design yields a clear event-study profile around the first verified strike. Figure 3 reports the event-study coefficients of equation (1) at the 5 km buffer, separately for the baseline panel (verified direct hits) and the broad panel (any attributed strike). The pre-strike profile is flat in the baseline panel: the coefficients before the first strike are small, centered close to zero, and jointly insignificant (joint $p = 0.98$). This pattern provides no evidence that refineries about to be struck were already on differential radiance trajectories.

After the first verified strike, radiance falls sharply and remains below the pre-strike path throughout the observed event window. The decline appears in the month of the first strike, deepens in the following months, and does not visibly reverse within the horizons

observed in the figure. The broad panel produces a noisier but similar profile. Taken together, the event-study evidence motivates summarizing the post-entry period with an aggregate estimate while also showing the dynamic pattern behind that aggregate.

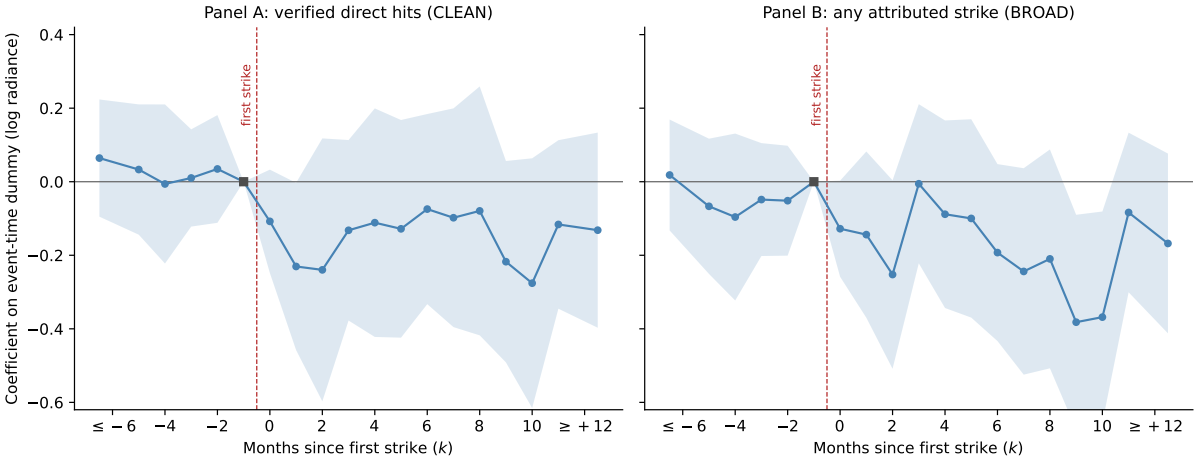


Figure 3: Event study: log radiance around the first verified strike (5 km buffer)

Notes. Event-study coefficients from equation (1): log monthly mean of quality-screened daily Black Marble radiance within 5 km of the refinery centroid (inner 500 m disk excluded), refinery and year-month fixed effects, reference period $k = -1$, endpoints binned at $k \leq -6$ and $k \geq +12$. Shaded bands are 95% confidence intervals from standard errors clustered at the refinery level. Panel A: baseline panel (22 treated, 7 never struck). Panel B: broad panel (24 treated, 8 never struck). *Source:* authors' calculations from NASA Black Marble and the verified ACLED strike record.

Table 1 reports the aggregate post-treatment effect across cumulative buffers from 1 km to 25 km; Appendix F shows the same gradient graphically. The decline is largest close to the refinery and attenuates monotonically with distance. In the baseline panel, radiance falls by 29.8% in the 1 km buffer, 22.9% within 2 km, 17.4% within 5 km, and 13.8% within 10 km. At 25 km the point estimate is smaller, 7.3%, and only marginally significant. The broad panel yields a similar profile, with a somewhat smaller estimate at 10 km and an essentially null estimate at 25 km. This monotonic attenuation is consistent with local industrial disruption around the refinery rather than a broad regional shock.

The 5 km buffer is our primary scale, broad enough to capture the refinery and its immediate industrial perimeter yet local enough to separate facility-level disruption from regional shocks. At this scale the decline is significant at the 5% level and, given the modest number of refinery clusters, survives wild cluster bootstrap-t inference; it is of similar size when the first one or two post-strike months are excluded, and the broad panel yields a somewhat smaller but still significant decline.

Table 2 decomposes the monthly effect jointly by event-time bin and spatial zone in the baseline panel. The decline appears immediately after the first verified strike and persists: at the 5 km buffer it is present in every event-time bin, with no evidence of recovery within the observed window. The estimates are less precise at longer horizons, because the last bin is identified only by refineries struck early enough in the campaign to be followed for

Table 1: Monthly DiD: aggregate post-treatment effect by distance from refinery

	1 km	2 km	5 km	10 km	25 km
Panel A: baseline panel (verified direct hits; 22 treated + 7 never struck)					
β_{post}	-0.354***	-0.261***	-0.192**	-0.149**	-0.076*
	(0.112)	(0.090)	(0.076)	(0.061)	(0.042)
Implied % change	−29.8	−22.9	−17.4	−13.8	−7.3
Observations	1,266	1,271	1,279	1,293	1,303
Panel B: broad panel (any attributed strike; 24 treated + 8 never struck)					
β_{post}	-0.375***	-0.274***	-0.167**	-0.102	-0.025
	(0.118)	(0.091)	(0.071)	(0.065)	(0.045)
Implied % change	−31.2	−24.0	−15.4	−9.7	−2.5
Observations	1,401	1,406	1,415	1,431	1,441

Notes. Aggregate post-treatment coefficient β_{post} from equation (2). Dependent variable: log monthly mean of quality-screened daily Black Marble radiance within the indicated cumulative buffer around the refinery centroid; every buffer excludes the inner 500 m disk. Refinery and year-month fixed effects. Standard errors clustered at the refinery level in parentheses. Implied % change is $\exp(\beta) - 1$. Wild cluster bootstrap-t p-values for the 5 km baseline estimate: 0.012–0.013. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

more than a year.

The spatial decomposition reinforces the local interpretation. The nearest measured zone, the $(0.5, 1]$ km ring, shows large declines in every post-treatment bin. The cumulative 2 km and 5 km buffers show persistent negative effects as well. The contraction remains significant in the next ring out: the $(1, 2]$ km ring falls by about 17.5% over the full post-treatment period, comparable to the 5 km buffer, though less stable across individual event-time bins. Only beyond 2 km do the estimates weaken: the $(2, 5]$ and $(5, 10]$ km rings are mostly negative but imprecisely estimated. The isolated significant estimate in the $(10, 25]$ km ring in months 4–6 is not mirrored in adjacent bins and is not interpreted as evidence of a separate discontinuous spillover at that distance.

The baseline estimate is stable across several diagnostics reported in Appendix G. Under the Rambachan–Roth relative-magnitudes framework, the aggregate effect in the baseline panel remains bounded away from zero for post-period violations up to $\bar{M} = 0.68$ times the largest pre-period deviation, and up to 1.3 times the median pre-period deviation. In the broad panel, whose pre-period is noisier, the corresponding breakdown values are smaller ($\bar{M} = 0.29$ and 0.53).

The estimate is also not driven by a single facility or cohort. Dropping one treated refinery at a time yields aggregate effects between −14.8% and −21.6%, all statistically significant at least at the 5% level; dropping one treatment cohort at a time leads to the same conclusion (Appendix G.2). Alternative measurement and functional-form choices produce similar magnitudes (Appendix G.3). Weighting the monthly aggregation by pixel counts

Table 2: Monthly DiD: implied percentage change by event-time bin and distance

	$[0,3]$	$[4,6]$	$[7,12]$	$[13,+)$	$[0,+)$
Panel A: cumulative buffers and nearest ring					
$(0.5,1]$ km ring	-28.5^{***}	-26.7^{***}	-29.4^{**}	-36.1^{***}	-29.8^{***}
Cumulative 2 km	-23.4^{***}	-20.3^{**}	-22.7^{**}	-25.5^{**}	-22.9^{***}
Cumulative 5 km	-20.0^{***}	-14.2	-18.4^{**}	-15.1	-17.4^{**}
Cumulative 10 km	-14.3^{***}	-13.7	-13.5^{*}	-13.7	-13.8^{**}
Cumulative 25 km	-6.3	-12.3	-4.0	-7.9	-7.3^{*}
Panel B: non-overlapping rings (derived)					
$(1,2]$ km	-11.3	-20.5^{**}	-20.7^{*}	-19.0	-17.5^{**}
$(2,5]$ km	-5.3	-4.6	-7.6	$+0.2$	-4.8
$(5,10]$ km	-9.0^{**}	-4.3	-7.5	$+2.6$	-5.5
$(10,25]$ km	-7.1^{*}	-13.9	-1.4	-5.9	-6.8

Notes. Baseline panel. Entries are implied percentage changes, $\exp(\beta)-1$, from a single regression per spatial zone in which the post-treatment indicator of equation (2) is replaced by indicators for event-time bins $[0,3]$, $[4,6]$, $[7,12]$, and $[13,+)$ months since the first verified strike; the $[0,+)$ column reports the single-indicator aggregate for reference. Refinery and year-month fixed effects; standard errors clustered at the refinery level. Panel B rings are constructed from pixel-count-weighted differences of the cumulative buffers. The $[13,+)$ bin is identified by the cohorts struck early in the campaign. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

or by valid observation days yields estimates of -14.6% and -15.4%, respectively, both significant at the 5% level. Poisson pseudo-maximum likelihood on radiance levels yields an estimate of -21.7%, significant at the 1% level. Across these variants, the implied local contraction remains on the order of 15–18%. Appendix J reports a descriptive repeated-strike intensity diagnostic: the contraction is already present among once-struck refineries, and lagged cumulative strike counts add no stable within-treated gradient once post-entry timing is controlled for.

5.2 Strike-day dynamics: daily IV

The daily IV analysis uses directional wind alignment to study strike-day dynamics at higher frequency. Table 3 reports the first stage on the post-first-strike estimation sample. The continuous alignment measure $Z_{it}=\cos(\phi_{it}-\theta_i)$ predicts strike incidence: the first-stage F-statistic is 12.0 in the baseline panel and 10.1 in the broad panel. The sign pattern reinforces the operational interpretation through a symmetry check: a binary headwind indicator predicts strike incidence with a negative coefficient of similar magnitude to the positive coefficient on the continuous alignment measure, as expected if winds opposed to the target direction reduce mission feasibility. A binary tailwind indicator, by contrast, is a weaker instrument than the continuous measure, consistent with feasibility varying smoothly with the angle between wind direction and the approach bearing. The results

are also robust to where the approach bearing is anchored: using late-2022 Ukrainian-controlled territory rather than the January 2022 international border leaves the first-stage and reduced-form patterns essentially unchanged (Appendix D).

Table 3: Daily IV: first stage and directional symmetry

	Baseline panel	Broad panel
Directional wind alignment Z_{it}	+0.0027***	+0.0028***
	(0.0008)	(0.0009)
Tailwind indicator ($\lvert\delta\rvert < 45°$)	+0.0018	+0.0017
	(0.0017)	(0.0017)
Headwind indicator ($\lvert\delta\rvert > 135°$)	-0.0027**	-0.0023*
	(0.0012)	(0.0012)
First-stage F (directional alignment)	11.96	10.14
Observations (refinery-days)	12,961	14,314
Strike-days	94	129

Notes. Each cell reports a separate regression of the binary strike-day indicator on the indicated wind variable, on the post-first-strike estimation sample, with refinery, month-of-year, year, and day-of-week fixed effects. Standard errors clustered at the refinery level in parentheses. The headwind row is a sign-symmetry check rather than a null placebo: if directional alignment drives mission feasibility, winds opposed to the approach bearing should reduce strike probability. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 4 plots the reduced-form coefficient of the horizon-h outcome on directional wind alignment. Throughout the IV analysis, horizon h denotes the seven-day centered window around day $t + h$, as defined in Section 4.2. Table 4 reports selected horizons together with Anderson–Rubin confidence intervals for the corresponding 2SLS coefficients. The reduced-form profile is consistent with a fire-then-damage dynamic. Radiance increases at short horizons: by 2.4% at $h = 0$, significant at the 10% level, and by 3.1% at $h = 14$, significant at the 1% level in the baseline panel. We interpret this initial increase as the combined light signature of active combustion, emergency response, and short-run operational disturbance around the facility. The interpretation is supported by independent FIRMS active-fire detections around recorded strike dates, discussed in Section 6.2.

The sign turns negative at longer horizons. At $h = 180$, the reduced-form coefficient is approximately -4.4% in the baseline panel and -4.0% in the broad panel, significant at the 1% level. The negative tail remains visible at $h = 270$, although it is estimated less precisely in the baseline panel. This timing is informative. The short-run positive response is consistent with the immediate fire and emergency-response component of a strike; the medium-run decline is consistent with a subsequent loss of ordinary nighttime illumination around damaged or partially disrupted refinery infrastructure.

The Anderson–Rubin intervals reinforce the same interpretation without requiring strong first-stage assumptions at every horizon. At $h = 14$, the interval is strictly positive in both

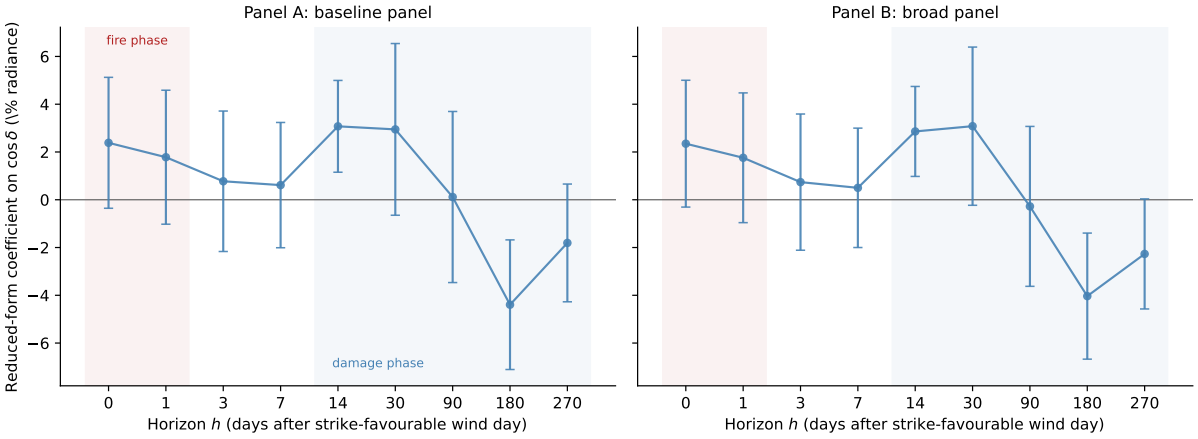


Figure 4: Daily IV: reduced-form response of radiance to directional wind alignment, by horizon

Notes. Reduced-form coefficient, expressed in percent of radiance, of the horizon-h outcome on directional wind alignment $Z_{it} = \cos(\phi_{it} - \theta_i)$. The outcome is the seven-day centered moving average of log radiance in the 5 km buffer around day $t + h$. The estimation sample is the post-first-strike refinery-day panel. All specifications include refinery, month-of-year, year, and day-of-week fixed effects. Vertical bars are 95% confidence intervals from standard errors clustered at the refinery level; weak-IV-robust inference is reported in Table 4. *Source:* authors' calculations from NASA Black Marble daily, ERA5-Land reanalysis, and the verified ACLED strike record.

Table 4: Daily IV: reduced form and weak-IV-robust inference, 5 km buffer

Horizon h (days)	RF (%)	AR 95% CI (2SLS)	$F(h)$	N
Panel A: baseline panel				
0	$+2.38^{*}$	$[-1.6, +\infty)$	5.3	10,847
14	$+3.07^{***}$	$[+3.2, +\infty)$	8.1	10,613
30	+2.95	$[-2.1, +36.4]$	9.9	10,339
90	+0.11	$[-19.5, +24.6]$	6.3	9,452
180	-4.39^{***}	$[-25.1, -3.5]$	14.6	7,896
270	−1.81	$[-33.6, +1.9]$	7.3	6,540
Panel B: broad panel				
0	$+2.35^{*}$	$[-1.1, +\infty)$	5.5	11,939
14	$+2.86^{***}$	$[+2.6, +\infty)$	6.3	11,683
30	$+3.08^{*}$	$[-0.6, +39.5]$	9.2	11,368
90	−0.28	$[-\infty, +\infty)$	3.7	10,410
180	-4.03^{***}	$[-25.3, -2.3]$	12.4	8,764
270	-2.27^{*}	$[-23.4, +0.1]$	10.1	7,326

Notes. 5 km cumulative buffer. The outcome at each horizon is the seven-day centered moving average of log radiance around day $t + h$, and horizon labels denote the center of the window. RF is the reduced-form coefficient on directional wind alignment $Z_{it} = \cos(\phi_{it} - \theta_i)$, expressed in percent; stars denote significance from standard errors clustered at the refinery level. The Anderson–Rubin 95% confidence interval refers to the coefficient on the binary strike-day indicator instrumented by Z_{it}, in log-radiance units, and provides weak-instrument-robust inference. $F(h)$ is the horizon-specific first-stage F-statistic on the estimation sample at horizon h. The 2SLS estimate at the damage horizon $h = 180$ is reported in Table 5, and the full set of 2SLS point estimates and standard errors in Appendix G.7. All specifications include refinery, month-of-year, year, and day-of-week fixed effects. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

panels. At $h = 180$, it is strictly negative in both panels: $[-25.1, -3.5]$ in the baseline panel and $[-25.3, -2.3]$ in the broad panel. At the transition horizons, especially between $h = 30$ and $h = 90$, the intervals are wide or unbounded, and we do not attach a separate interpretation to those estimates.

The discrete horizons in Table 4 sample the response path. As a complementary check, we also estimate reduced forms in which the outcome is averaged over wider day bins covering the year after the potential strike day, together with a pre-event placebo bin. The placebo exercise shows no detectable association between wind alignment on day t and radiance over days $[-45, -10]$ before t ($+0.1\%$, $p = 0.92$). In the post-period, radiance is higher over days $(15, 90]$ and lower over days $(91, 180]$ and $(181, 270]$. This binned evidence supports the sign reversal in the discrete-horizon estimates while showing that the positive phase extends beyond the active-fire window. This pattern is consistent with an elevated-light phase associated with the strike and its immediate operational aftermath, rather than with fires themselves persisting for several months. Appendix G.6.

The 2SLS coefficients rescale these reduced forms by a first stage that captures the wind-sensitive margin of strike incidence. Instrument strength varies across horizons with outcome availability: the pooled first stage clears the conventional rule of thumb ($F = 12.0$ in the baseline panel and $F = 10.1$ in the broad panel), and at the damage horizon $h = 180$ the horizon-specific first stage is comparably strong ($F = 14.6$ and $F = 12.4$, respectively), whereas it is weaker at the short and transition horizons. We therefore use the Anderson–Rubin intervals in Table 4 as the preferred weak-instrument-robust inference at those horizons.

Table 5: Daily IV: first and second stage at the damage horizon ($h = 180$, 5 km buffer)

	Baseline panel	Broad panel
First stage. Dependent variable: strike-day indicator		
Directional wind alignment Z_{it}	$+0.0043^{***}$	$+0.0048^{***}$
	(0.0011)	(0.0014)
First-stage F	14.6	12.4
Second stage. Dependent variable: seven-day mean log radiance at $h = 180$		
Strike-day indicator (instrumented)	-10.31^{***}	-8.47^{***}
	(3.25)	(2.83)
AR 95% CI	$[-25.1, -3.5]$	$[-25.3, -2.3]$
Observations	7,896	8,764

Notes. Daily IV estimated at horizon $h = 180$ days on the post-first-strike sample, 5 km buffer. The outcome is the seven-day centered moving average of log radiance around day $t + 180$. The first stage regresses the binary strike-day indicator on directional wind alignment $Z_{it} = \cos(\phi_{it} - \theta_i)$. The second stage instruments the strike-day indicator with Z_{it}; coefficients are reported in log-radiance units. Both stages are estimated on the $h = 180$ sample, so the first-stage coefficient differs from the pooled first stage in Table 3. Standard errors clustered at the refinery level in parentheses. F is the horizon-specific first-stage F-statistic. The Anderson–Rubin interval provides weak-instrument-robust inference for the second-stage coefficient. Because the 2SLS coefficient rescales the reduced form by the wind-sensitive change in daily strike probability, its magnitude should not be read as the average effect of a typical strike. Estimates at all horizons appear in Table 4 (reduced form and AR intervals) and Appendix G.7 (full 2SLS grid). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5 reports the 2SLS estimates at $h = 180$, with the full grid across horizons in Appendix G.7. The coefficients are large and negative, and the Anderson–Rubin intervals exclude zero in both panels; their magnitude reflects the narrow wind-sensitive margin of daily strike incidence among refineries already exposed to repeated attacks, and should not be read as the average effect of a typical strike or compared mechanically with the campaign-level monthly DiD estimate. For interpretation, the key object remains the dynamic profile: strike-favorable wind conditions predict brighter nights at short horizons and darker nights around the same facilities at medium horizons.

5.3 Target heterogeneity

A possible alternative interpretation of the headline result is that the radiance contraction around struck refineries reflects generic disruption associated with high-intensity strike activity: air defense activation, precautionary shutdowns, local restrictions, evacuations, or temporary dimming. The deep-territory strike record allows us to assess this interpretation by comparing refineries with other classes of targets struck by Ukraine inside Russia during the same period. If the estimated decline mainly reflected a generic local footprint of strikes, rather than disruption to productive infrastructure, we would expect similar radiance losses around other struck targets.

We estimate the aggregate post-treatment specification in equation (2) for 106 non-refinery strike sites drawn from the same verified ACLED extract and classified into five target categories: military targets (46 sites), fuel depots (10), other industrial plants (22), power infrastructure (17), and transport nodes (11). Each site is dated by its first recorded strike, and the coefficient is the average effect over all months following that strike, as in the refinery specification. Identification within each non-refinery category comes from the staggered timing of first strikes among sites in that category. The exercise uses the same satellite product, quality screening, monthly aggregation, 5 km buffer, and fixed-effects structure used for the refinery estimates.

Table 6: Target heterogeneity: aggregate post-strike effect by target category

Target category	β_{post}	(SE)	Implied %	Sites	Pre-trend p-value
Oil refineries (baseline panel)	-0.192**	(0.076)	−17.4	29	0.83
Military targets	-0.017	(0.034)	−1.7	46	0.15
Fuel depots	+0.082	(0.108)	+8.6	10	0.96
Other industrial plants	-0.040	(0.041)	−3.9	22	0.37
Power infrastructure	-0.021	(0.096)	−2.1	17	0.35
Transport nodes	+0.012	(0.107)	+1.2	11	< 0.001

Notes. Aggregate post-treatment coefficient from equation (2), estimated separately by target category, equal to the average effect over all months following the first strike. The dependent variable is log monthly mean of quality-screened daily Black Marble radiance within 5 km of the site; the inner 500 m disk is excluded. All specifications include unit and year-month fixed effects. Standard errors clustered at the unit level in parentheses. The refinery row reports the baseline-panel headline estimate, with 22 treated and 7 never-struck refineries. Non-refinery categories contain treated sites only and are identified by staggered first-strike timing within each category. The pre-trend column reports the p-value of the joint test that the event-time coefficients $k \in [-6, -2]$ are zero. The transport category fails this diagnostic and its estimate should not be interpreted. The same comparison at the 1 km buffer, where the refinery effect is sharpest, is reported in Appendix H. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6 shows that the refinery effect is not replicated across other target categories. Military targets, the largest comparison group, have a small and statistically insignificant estimate of −1.7%. Other industrial plants, power infrastructure, and transport nodes also show no statistically significant contraction, although the transport category fails the pre-trend diagnostic and is not interpreted. Fuel depots have a positive and imprecisely estimated coefficient. This pattern is useful because fuel depots are often associated with visible fires but do not anchor the same surrounding production process as refineries.

The contrast remains visible at the 1 km buffer. The refinery estimate rises to −29.8%, while no other category develops a statistically significant decline (Appendix H). The comparison therefore weakens the interpretation that the headline result reflects a generic local effect of strikes, air defense activity, or wartime disruption around targeted sites.

The evidence is instead consistent with a more specific interpretation: the radiance loss is concentrated around targets that combine physical damage with dense, continuous, and

spatially fixed productive activity. Refineries are large continuous-process facilities, and the loss points to disruption of the surrounding activity they anchor—labor, logistics, storage, maintenance, and auxiliary industrial operations—rather than to a generic response to being struck.

6 Spatial evidence and validation

This section interprets the radiance response through two lenses: how the daily IV response is distributed in space, and an independent satellite record of active fires from NASA FIRMS. The aim is to distinguish a transient, fire-related light signature from a more persistent loss of ordinary nighttime illumination around the refinery.

6.1 Spatial signature of fire and damage

Section 5.2 characterized the daily IV response over time at the 5 km buffer. Here we instead examine how that response is distributed in space, re-estimating the reduced form at each cumulative buffer and ring. The short-run increase and the medium-run decline differ in their spatial pattern, and this contrast is informative. The short-run increase is visible at aggregated buffer scales: at $h = 14$ days, the reduced-form coefficient on directional wind alignment is positive and significant at the 5, 10, and 25 km cumulative buffers (+3.1%, +2.7%, and +2.3%), while in the 1 and 2 km buffers the point estimates are positive but less precisely estimated.

Because the pixels closest to the facility account for a large share of the nighttime light within a cumulative buffer, a localized increase in radiance at or near the facility raises mean radiance in larger cumulative buffers as well; the signal at wider buffers therefore does not imply a spatially diffuse fire. At the same time, the smallest buffers are more exposed to missingness and quality screening in the daily Black Marble product, especially around active fires, smoke, clouds, and other short-run disturbances. The short-run IV response should therefore be read as a transient fire-related light signature, measured with noise at the finest spatial scales.

The medium-run decline is more informative about damage to ordinary illumination. At $h = 180$ days, the reduced-form coefficient in the nearest measured ring, $(0.5, 1]$ km from the refinery centroid, is -4.1% and statistically significant at the 5% level, close to the -4.4% estimate in the cumulative 5 km buffer. Coefficients remain negative at wider cumulative buffers, but cumulative buffers are dominated by the pixels closest to the facility and therefore do not by themselves establish that the damage is spatially diffuse. Daily ring-level estimates at long horizons are imprecise because constructing daily ring outcomes requires valid pixels in adjacent buffers simultaneously, reducing the usable sample. The more reliable spatial evidence on the location of the contraction therefore

comes from the monthly design: the ring decomposition in Table 2 places the largest and most persistent losses within the inner two kilometers of the facility, with much weaker evidence of contraction in the $(2, 5]$ and $(5, 10]$ km rings.

6.2 Independent satellite validation: NASA FIRMS

NASA FIRMS provides an independent satellite-based check on the physical fire component of the verified strike record. Figure 5 reports the share of verified direct hits with at least one FIRMS thermal-anomaly detection within 5 km of the refinery, by day relative to the verified strike date. The pre-strike baseline is 3.9%, computed over days −14 to −4. On the recorded strike date, the detection rate rises to 7.9%, roughly twice the baseline. The elevation is concentrated around the strike date, allowing for the fact that ACLED event dates and satellite acquisition dates can fall on opposite sides of midnight.

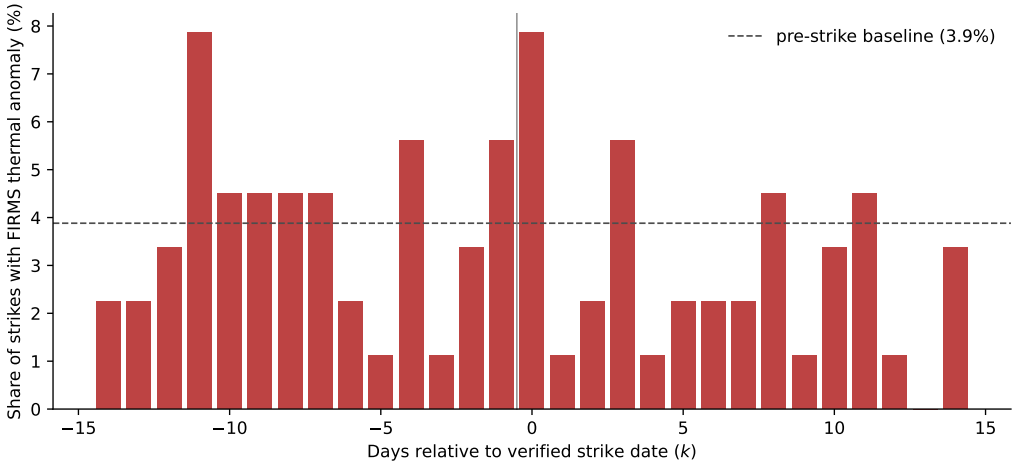


Figure 5: FIRMS thermal anomalies around verified strike dates

Notes. Share of verified direct hits with at least one FIRMS thermal-anomaly detection within 5 km of the refinery, by day relative to the verified strike date. The dashed horizontal line is the pre-strike baseline, computed as the mean over days −14 to −4. Sample: 89 verified direct hits within FIRMS coverage. *Source:* authors' calculations from NASA FIRMS and the verified ACLED strike record.

This validation is deliberately conservative. FIRMS detects only a subset of physical strike events: detection depends on satellite overpass timing, cloud cover, the duration and intensity of combustion, and whether the strike produces a thermal signature large enough to be observed from space. In addition, refineries can exhibit routine thermal activity unrelated to attacks, including flaring. For these reasons, the FIRMS exercise is not treated as a census of strikes. It is used instead as an independent check that a measurable fire signal appears at the recorded strike date for a nontrivial share of verified direct hits.

The FIRMS evidence is therefore complementary to the IV. It supports the interpretation of the short-run positive radiance response as a fire-related light signature rather than

as an arbitrary measurement artifact. The subsequent decline in nighttime radiance is interpreted separately, through its timing and spatial concentration: it appears at medium horizons, is strongest close to the refinery, and aligns with the monthly DiD evidence on persistent local contraction.

6.3 Cross-strategy convergence

The two strategies estimate different parameters. The monthly DiD identifies the cumulative effect of entry into the strike campaign, which includes persistent damage, repeated strikes, operational adjustment, and the local consequences of disrupted refinery activity. The daily IV identifies the dynamic response to the component of strike-day variation shifted by wind conditions, among refineries already exposed to the campaign. A strict numerical reconciliation between the two estimates is therefore not warranted.

The convergence across designs is nevertheless informative. The sign is consistent: both strategies point to a loss of local radiance after refineries are struck. The timing is consistent: the monthly DiD shows a persistent decline after campaign entry, while the daily IV separates an initial positive light signature from a delayed negative response. The two designs also agree on spatial scale: the monthly ring decomposition and the spatial reading of the daily IV both place the largest contractions close to the facility rather than at the regional scale, and the short-run fire component is corroborated by the independent FIRMS detections reported above. This convergence comes from two distinct sources of variation applied to the same quality-screened satellite product, so the interpretation does not rest on any single specification.

7 Discussion

The estimates imply a large local contraction in satellite-visible activity around struck refineries. In the monthly DiD, radiance falls by roughly 15–18% within 5 km of the refinery and by about 30% in the nearest measured ring. These magnitudes are large by the standards of the nightlights literature, but the spatial pattern helps discipline their interpretation. The decline is strongest close to the facility, attenuates with distance, and is no longer detectable at 25 km. This pattern points to disruption in the refinery and its immediate industrial perimeter rather than to a broad contraction of the surrounding region.

To provide an order-of-magnitude interpretation, we apply standard lights-to-GDP elasticities from the nightlights literature to the radiance coefficients, without converting the estimates into currency units. Using a lights elasticity of $\psi \simeq 0.30$, close to the benchmark estimates in Henderson et al. (2012), a radiance decline of 0.16–0.20 log points within 5 km corresponds mechanically to a local activity shortfall of about 5–6%. The 0.35 log-point

decline in the nearest measured ring corresponds to roughly 11%. These figures are scale anchors rather than estimates of local GDP. The elasticity is estimated from aggregate relationships over longer horizons, whereas our estimates refer to small buffers around highly capital-intensive industrial facilities. In this setting, a given radiance decline may correspond to a larger loss of capital utilization and a smaller employment response than in a typical urban area. The exercise is therefore useful for judging economic order of magnitude, not for valuing the damage.

The contraction is also persistent over the horizons observed in the panel. The decline appears in the month of the first verified strike and shows no evidence of reverting within the event window. In the nearest measured ring, the estimated decline remains large beyond the thirteenth month, although that long-horizon bin is identified by the cohorts struck early enough in the campaign to be followed for more than a year. Two non-exclusive interpretations are consistent with this persistence. First, repairs to damaged refining infrastructure may be slow, especially when replacement equipment, specialized components, or technical services are constrained by sanctions and wartime logistics. Second, the treatment is not a single isolated event. Refineries that enter the campaign often experience further strikes, so the observed post-treatment path reflects a contest between repair, operational adaptation, and renewed damage. Whether the contraction would attenuate if the strikes stopped is not identified in our panel. What the data show is that, under the campaign as conducted, the affected industrial perimeters do not return to their pre-strike trajectory over the horizons we observe.

Several economic channels could generate a contraction concentrated near the facility, although the data do not allow us to distinguish them directly. The most immediate channel is reduced operation at damaged refining units. If a distillation column, cracking unit, storage facility, or auxiliary system is taken offline, the loss of ordinary nighttime illumination may come from the affected unit itself, from lower activity in adjacent loading, storage, maintenance, and logistics areas, or from temporary reorganization inside the refinery complex. A related channel is reduced activity among firms and workers that serve the refinery in the surrounding industrial area, including transport and logistics contractors, equipment-maintenance firms, and chemical-service suppliers. A more diffuse channel would run through nearby settlements serving the refinery workforce. The spatial evidence is more consistent with the first two types of channel than with a broad regional demand effect: the largest declines occur close to the facility and in the inner industrial zone, not at the regional scale.

Two considerations delimit the scope of the estimates. First, nighttime radiance is a proxy for local economic activity rather than a direct measure of output. The elasticity-based scaling above provides a sense of magnitude, not a measurement of local GDP (Henderson et al., 2012; Beyer et al., 2022). A possible alternative interpretation is that the post-strike

decline reflects a deliberate darkening of the facility rather than a genuine reduction in activity: operators might cut lighting once a refinery comes within reach of Ukrainian long-range drones, in the hope of presenting a harder target. Two features of the setting make this unlikely to account for the estimates. First, contemporary long-range strike drones typically navigate to preset geographic coordinates using inertial and satellite guidance rather than by visually acquiring an illuminated target; for a fixed industrial site whose coordinates are already known, switching off lights offers little protection. Second, the contraction is not confined to the facility perimeter. It extends through the inner two kilometers and attenuates only gradually with distance, whereas a darkening intended to protect the refinery would be concentrated on the installation itself. Deliberately blacking out the surrounding settlements over several square kilometers would not serve as a point-defense measure, while a genuine reduction in refinery operations, and in the activity of the firms and workers in its immediate industrial perimeter, produces the spatial pattern we observe. A deliberate blackout of the installation could in principle contribute to the innermost ring, but it cannot explain a contraction that remains visible well beyond it.

Second, the estimates identify the local radiance footprint of disruption around struck facilities; they do not identify general-equilibrium adjustments such as the reallocation of refining runs to undamaged plants, changes in imports, inventories, fuel prices, or fiscal revenue. Quantifying the aggregate cost of the campaign would require combining these local estimates with production, trade, fiscal, and price data.

The setting also limits the temporal reach of the conclusions. Wartime reporting from inside Russia may miss some successful strikes. To the extent that such omissions assign treated periods to nominally untreated or not-yet-treated refinery-months, they would attenuate the estimated contraction. At the same time, the campaign is ongoing. The persistence documented here extends through and beyond the first post-strike year for early-treated refineries, but longer-run trajectories—recovery, adaptation, or further decline—will become identifiable only as the panel lengthens.

These limits notwithstanding, the exercise is valuable precisely because of the setting it studies. The local economic consequences of the refinery campaign unfold inside a wartime authoritarian economy, where official information is limited, strategic infrastructure losses are politically sensitive, and reported damage is often downplayed or denied. This opacity makes it difficult to answer basic policy-relevant questions: whether deep strikes produce measurable economic effects, whether affected refineries experience persistent reductions in local activity, how large those losses are at the facility perimeter, and whether such effects can be credibly measured absent reliable production data. Our approach provides externally verifiable evidence on these margins. By combining a manually verified strike record, quality-screened satellite radiance, independent fire detections, and weather-based

variation in strike incidence, we show that refinery strikes are followed by a persistent local contraction, locate that contraction in space, and separate the immediate fire signature from the subsequent decline in ordinary nighttime illumination. The analysis does not settle broader questions about the aggregate cost or strategic effectiveness of the campaign, but it shows that those questions can be disciplined with evidence even in a setting designed to make economic losses hard to observe.

8 Conclusion

The evidence points to a coherent empirical pattern. Ukrainian strikes on Russian refineries generate a short-run, fire-related increase in nighttime radiance, followed by a persistent decline in ordinary nighttime illumination around the affected facilities. The monthly DiD shows that entry into the refinery-strike campaign is followed by a large and spatially ordered contraction: radiance falls by roughly 15–18% within 5 km of the refinery and by about 30% in the nearest measured ring, and is no longer detectable at 25 km. The daily IV, which exploits wind-driven variation in strike incidence, separates the immediate light signature of a strike from the subsequent decline at medium horizons. Independent FIRMS detections validate the physical fire component, while the spatial gradient places the persistent contraction close to the refinery and its industrial perimeter.

The two empirical designs do not estimate the same parameter, and their magnitudes should not be mechanically reconciled. The monthly design captures the cumulative effect of entering the strike campaign, including repeated strikes, repair, operational adjustment, and renewed damage. The daily IV captures the response to wind-sensitive strike-days among refineries already exposed to the campaign. Their value lies in their complementarity: different sources of variation, at different frequencies, produce a consistent account of fire, damage, persistence, and spatial concentration. The comparison with other deep-territory targets reinforces this interpretation, since comparable contractions do not appear around other categories of struck sites when measured with the same satellite product and empirical specification.

These findings speak beyond the specific campaign studied here. They show that fossil-fuel refining infrastructure can function as a critical node of local economic activity, and that disrupting such fixed industrial assets can generate persistent losses in the surrounding industrial economy. In contemporary conflicts, relatively low-cost long-range strike systems make productive infrastructure vulnerable far from the front line. The results therefore connect the economics of conflict to broader questions in political economy and applied economics: how shocks to critical industrial infrastructure propagate through local economies, and how those effects can be measured when the relevant data are unavailable, delayed, or politically opaque.

References

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1):113–132.

Anderson, M. L. (2020). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*, 18(4):1886–1927.

Anderson, T. W. and Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, 20(1):46–63.

Beyer, R. C. M., Hu, Y., and Yao, J. (2022). Measuring quarterly economic growth from outer space. IMF Working Paper 2022/109, International Monetary Fund.

Dabrowski, M. (2025). How resilient is Russia's economy after four years of war? Working Paper 32/2025, Bruegel.

Daianu, D., Grosu, T., Neacsu, A., Tanase, A., and Vranceanu, R. (2025). Macroeconomic consequences of the war in Ukraine on central and eastern European economies: A SVAR analysis. *Defence and Peace Economics*, pages 1–25. doi: 10.1080/10242694.2025.2569771.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12):4178–4221.

Devadoss, S. and Ridley, W. (2024). Impacts of the Russian invasion of Ukraine on the global wheat market. *World Development*, 173:106396.

Donaldson, D. and Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–198.

Gopinath, G., Gourinchas, P.-O., Presbitero, A. F., and Topalova, P. (2025). Changing global linkages: A new cold war? *Journal of International Economics*, 153:104042.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Herrnstadt, E., Heyes, A., Muehlegger, E., and Saberian, S. (2021). Air pollution and criminal activity: Microgeographic evidence from Chicago. *American Economic Journal: Applied Economics*, 13(4):70–100.

Kaushal, S. and Suess, J. (2025). Net assessment of Russian and NATO capabilities in a modern strike campaign. Occasional paper, Royal United Services Institute.

Korovkin, V. and Makarin, A. (2023). Conflict and inter-group trade: Evidence from the 2014 Russia-Ukraine crisis. *American Economic Review*, 113(1):34–70.

Korovkin, V., Makarin, A., and Miyauchi, Y. (2025). Supply chain disruption and reorganization: Theory and evidence from Ukraine's war. *Review of Economic Studies*, page rdaf080. doi: 10.1093/restud/rdaf080.

KSE Institute (2025). Russia chartbook: Budget deficit rises sharply as oil and gas revenues drop, the economy stalls, and expenditures continue to soar. Technical report, Kyiv School of Economics Institute.

Luo, H., Liang, Z., Zhu, M., Hu, X., and Wang, G. (2018). Integrated optimization of unmanned aerial vehicle task allocation and path planning under steady wind. *PLOS ONE*, 13(3):e0194690.

Mahmood, R. and Jetter, M. (2023). Gone with the wind: The consequences of us drone strikes in Pakistan. *The Economic Journal*, 133(650):787–811.

Martínez, L. R. (2022). How much should we trust the dictator's GDP growth estimates? *Journal of Political Economy*, 130(10):2731–2769.

Martinez, S.-K., Pompeo, M., Sheremeta, R. M., Vakhitov, V., Weber, M., and Zaika, N. (2026). Civilian evacuation during war: Evidence from Ukraine. *The Economic Journal*, 136(675):1155–1171.

Mellon, J. (2025). Rain, rain, go away: 194 potential exclusion-restriction violations for studies using weather as an instrumental variable. *American Journal of Political Science*, 69(3):881–898.

Miguel, E. and Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, 96(1):1–15.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N. (2021). ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9):4349–4383.

Persico, C. and Venator, J. (2021). The effects of local industrial pollution on students and schools. *Journal of Human Resources*, 56(2):406–445.

Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591.

Riaño, J. F. and Valencia Caicedo, F. (2024). Collateral damage: The legacy of the secret war in Laos. *The Economic Journal*, 134(661):2101–2140.

Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., Schultz, L., Bell, J., Stokes, E. C., Pandey, B., Seto, K. C., Hall, D., Oda, T., Wolfe, R. E., Lin, G., Golpayegani, N., Devadiga, S., Davidson, C., Sarkar, S., Praderas, C., Schmaltz, J.,

Boller, R., Stevens, J., Ramos González, O. M., Padilla, E., Alonso, J., Detrés, Y., Armstrong, R., Miranda, J., Conte, Y., Marrero, N., MacManus, K., Esch, T., and Masuoka, E. J. (2018). NASA's Black Marble nighttime lights product suite. *Remote Sensing of Environment*, 210:113–143.

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A. (2014). The new VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, 143:85–96.

Stastny, T. and Siegwart, R. (2019). On flying backwards: Preventing run-away of small, low-speed, fixed-wing UAVs in strong winds. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5198–5205. IEEE.

Stockholm Institute of Transition Economics (2025). Financing the Russian war economy. Technical report, European Council.

Voskuijl, M., Dekkers, J., and Savelsberg, A. (2020). Flight performance analysis of the Samad family of attack drones. *Science & Global Security*, 28(3):113–134.

Yermakov, V. (2025). Fiscal flex: Russia's oil and gas revenues in 2024. Oxford energy comment, Oxford Institute for Energy Studies.

A Refinery list

Table 7 enumerates the operational Russian oil refineries in the panel (29 in the baseline panel, 32 in the broader variant), with coordinates, the date of the first manually verified direct hit, and verified event counts over the sample period (June 2022 – May 2026). Refineries are ordered by descending verified event count within the treated subset. The geographic distribution is visualized in Figure 6 below.

Table 7: Refineries in the panel

Refinery	Lat	Lon	First strike	Direct	All events
Treated — verified direct hit (22)					
Ryazan Oil	54.56	39.75	2024-03-13	12	15
Saratov Oil	51.45	45.94	2025-08-10	9	13
Volgograd	48.49	44.63	2024-02-03	8	13
Ilsky	44.86	38.62	2024-02-09	7	11
Syzran	53.09	48.39	2024-03-16	10	11
Kstovo (NORSI)	56.13	44.18	2024-03-12	6	10
Novokuibyshevsk	53.14	50.04	2024-03-16	6	10
Slavyansk	45.25	38.12	2024-03-17	7	9
Yaroslavl (YANOS)	57.55	39.81	2025-10-31	7	9
Afipsky	44.88	38.82	2023-10-29	6	8
Novoshakhtinsk	47.76	39.94	2022-06-22	6	7
Kirishi (KINEF)	59.48	32.05	2025-10-04	2	5
Orsk (Orsknefteorgsintez)	51.25	58.52	2025-10-03	3	4
Perm	57.92	56.13	2026-04-30	2	3
Moscow	55.64	37.80	2024-09-01	2	2
Ukhta	63.64	53.82	2025-08-10	1	2
Gazprom Neftekhim Salavat	53.40	55.90	2024-05-09	2	2
Taneko	55.64	51.82	2024-04-02	1	2
Tuapse	44.10	39.10	2023-02-27	1	1
Mariysky	56.98	47.73	2025-10-29	1	1
Novospassky (NS-Oil)	53.16	47.72	2025-10-29	1	1
Kuibyshev	53.10	50.05	2024-03-23	1	1
Treated in broad panel only — suspected/debris attribution (2)					
Krasnodar	45.00	38.97	2025-08-30	0	1
Novo-Ufa	54.82	56.05	2025-09-13	0	1
Never struck (8)					
JSC Antipinsky	57.10	65.76	—	0	0
Chernigov	55.61	86.14	—	0	0
Omsk	55.08	73.24	—	0	0
Angarsk Petrochemical	52.52	103.88	—	0	0
Achinsk	56.28	90.50	—	0	0
Komsomolsk	50.55	137.03	—	0	0
Khabarovsk	48.48	135.08	—	0	0
Ufimsky	54.85	56.10	—	0	0

Notes. Panel of 32 operational Russian oil refineries: 22 with at least one manually verified direct hit (baseline panel), 2 treated only under the broader attribution (suspected hits or falling debris), and 8 never struck. The "First strike" column reports the treatment date from the manually curated event-level verification of ACLED event descriptions (Section 3.3). "Direct" and "All events" count verified direct hits and all attributed events per refinery in the classified record; for co-located refinery clusters (Samara and Ufa areas), a small number of events reassigned in the curated treatment definition may be counted under the neighboring facility. All listed refineries enter the daily Black Marble estimation sample.

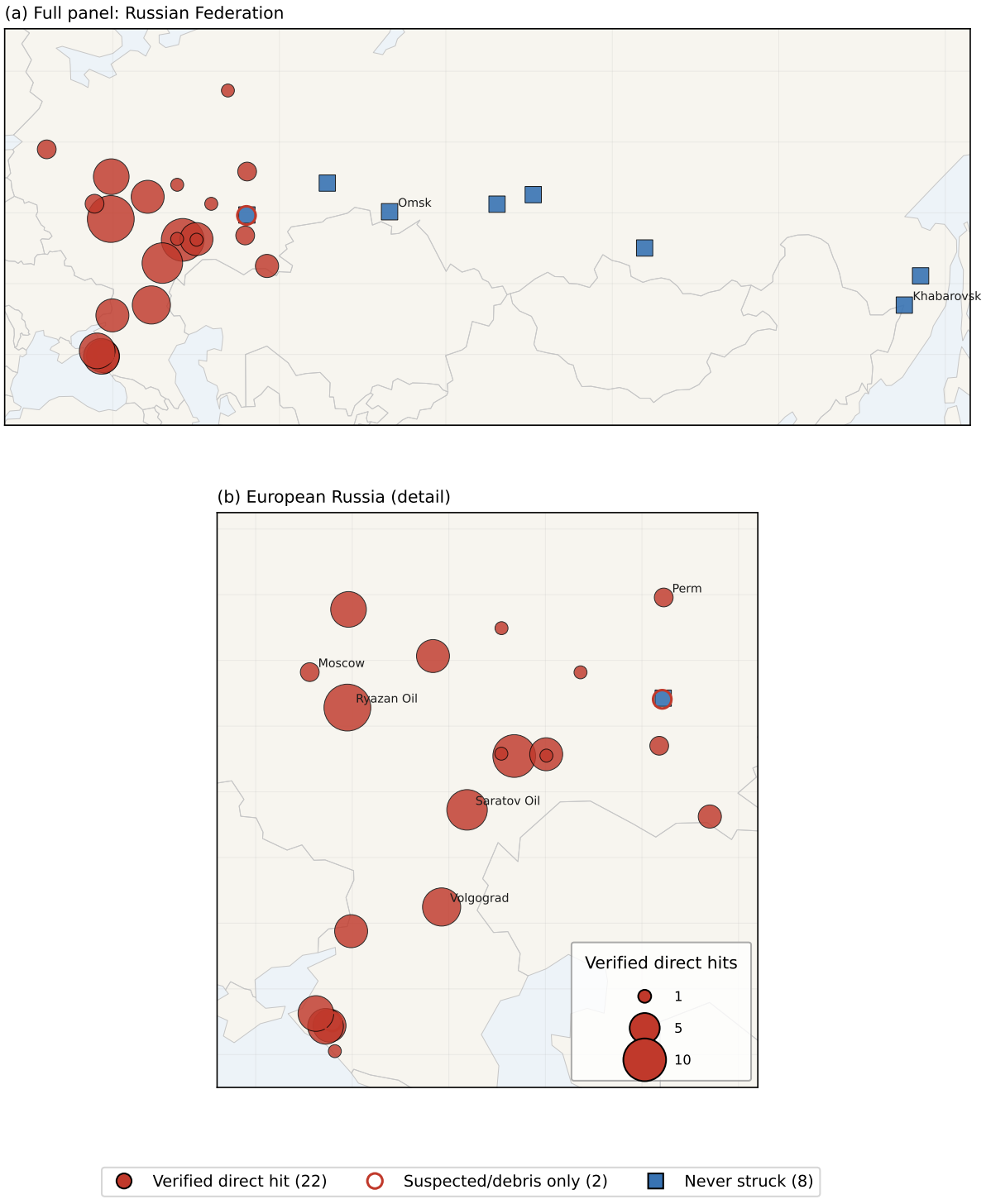


Figure 6: Geographic distribution of refineries in the analysis panels

Notes. Panel (a) shows the refineries used across the baseline and broad analysis panels; panel (b) zooms on European Russia, where struck refineries are concentrated. Filled red circles denote refineries with at least one verified direct hit. Circle size is proportional to the number of verified direct hits during the sample period. Open red circles denote refineries with only suspected-hit or debris-only attributions. Blue squares denote refineries never struck during the sample period. The baseline panel includes the 22 verified-direct-hit refineries and 7 never-struck controls; the broad panel adds the 2 suspected/debris-only refineries and one additional never-struck control. The headline estimate is robust to restricting the control group to western refineries only (Appendix G.4). Sample period: June 2022–May 2026. *Source:* authors' calculations from the rebuilt refinery panels and the verified ACLED strike record; country borders from Natural Earth.

B Construction and verification of the strike record

This appendix documents how the treatment record is built and verified, and reports the full accounting from raw candidate events to the analysis panels.

Discarding proximity matching. An earlier version of the record matched ACLED events to refineries by spatial proximity (events within 15 km of a refinery centroid). A manual audit showed that this proximity rule, rather than the underlying event data, was systematically unreliable: several refineries lie within a few kilometers of air bases, airfields, and other frequently struck installations, so the radius rule dates a refinery's first strike to attacks on those neighboring targets — the Saratov refinery to strikes on the Engels-2 air base, the Ryazan refinery to the Dyagilevo airfield, the Orsk refinery to an event near Voronezh. All proximity-based attributions were discarded.

Name-based attribution and manual classification. The record was rebuilt from the textual descriptions of the ACLED events. Candidate events are those whose description names a specific refinery or refinery complex (148 events). Each candidate was read and classified manually into three classes: *verified direct hit* (the description reports an impact, explosion, fire, or unit damage at the facility itself; 101 events), *suspected hit* (the facility is described as the suspected or possible target, without confirmed damage; 27 events), and *falling debris* (only debris from an intercepted drone fell on or near the facility, or the description reports no damage; 14 events). Six candidates were discarded as not referring to a strike on a panel refinery. Co-located clusters were disambiguated explicitly: the Samara cluster (Kuibyshev, Novokuibyshevsk, Syzran), the Ufa cluster (Novo-Ufa, Ufaneftekhim, Ufimsky), the Engels–Saratov pair (refinery versus the separate Kristall depot), and the Krasnodar area (the city refinery versus the Afipsky, Ilsky, and Slavyansk plants).

Constructing the analysis panels. Table 8 reports the accounting. Of the 33 refineries in the master list, Ufaneftekhim is excluded from both panels: two events are attributed to it but neither is a verified direct hit, so it can serve neither as a treated unit nor as a control. The remaining 32 units split into 22 refineries with at least one verified direct hit, 2 with suspected or debris attributions only (Krasnodar and Novo-Ufa), and 8 never-struck refineries. The baseline panel (29 units) consists of the 22 verified-direct-hit refineries and 7 never-struck controls; Ufimsky, in the Ufa cluster where suspected events cannot be attributed with confidence across co-located facilities, is conservatively excluded from the baseline and enters the broad panel only. The broad panel (32 units) treats any attributed event as treatment.

Table 8: Sample construction: from raw events to the analysis panels

Panel A: strike events	
Proximity-matched events (15 km rule), re-audited and discarded	284
Name-based candidate events (description names a refinery)	148
Discarded on manual reading (not a panel-refinery strike)	6
Attributed events	142
of which: verified direct hits	101
of which: suspected hits	27
of which: falling debris	14
Panel B: refineries	
Master list of operational refineries	33
Excluded (attributed events, none verified: Ufaneftekhim)	1
Units in the analysis panels	32
of which: verified direct hit (baseline treated)	22
of which: suspected/debris attribution only (broad treated)	2
of which: never struck	8
Baseline panel (22 treated + 7 never struck)[a]	29
Broad panel (24 treated + 8 never struck)	32

Notes. Event counts refer to the verified record spanning June 2022–May 2026 (Section 3.3).
[a] Ufimsky, in the Ufa cluster, is conservatively excluded from the baseline panel — suspected events in the cluster cannot be attributed with confidence across co-located facilities — and enters the broad panel as a never-struck unit.
In compliance with the ACLED Partner-Tier license, the event-level classification file with description excerpts is not redistributed; the counts above and the refinery-level variables we use are derived aggregations.

C Approach bearings for the daily IV

The daily IV described in Section 3.4 and Section 4.2 relies on a refinery-specific bearing θ_i from the nearest point on the January 2022 internationally recognized Ukraine–Russia border to each refinery. Table 9 reports the resulting bearings and the corresponding great-circle distances for the treated refineries in the panel. Because approach bearings are measured clockwise from due north and wrap around 360°, northern refineries fall on both sides of zero: Kirishi lies at $\sim$ 354°, just west of due north, while Yaroslavl, Ukhta, and Moscow have bearings in the north-northeast range (33–37°). Most Volga and Ural refineries fall in the northeast quadrant (45–72°), and the North Caucasus refineries on the southern flank (Slavyansk, Tuapse, Ilsky, Afipsky, and Krasnodar) lie around 149–165°. The panel therefore spans a wide range of approach bearings. Distances range from approximately 40 km (Novoshakhtinsk, just across the border) to roughly 1,700 km (Ukhta in the north), with the most distant Ural refineries—Perm and Orsk—around 1,400–1,500 km.

Table 9: Approach bearings θ_i and great-circle distances, treated refineries.

Refinery	Bearing θ_i (°)	Distance (km)
Kirishi (KINEF)	354	794
Yaroslavl (YANOS)	33	705
Ukhta	35	1,721
Moscow	37	463
Novoshakhtinsk	42	38
Perm	45	1,511
Ryazan	47	465
Kstovo (NORSI)	49	794
Taneko	49	1,157
Mariysky	51	1,029
Novospassky	51	786
Syzran	53	820
Novokuibyshevsk	56	921
Kuibyshev	56	920
Novo-Ufa	55	1,355
Gazprom Neftekhim Salavat	61	1,296
Saratov	60	589
Orsk	69	1,418
Volgograd	72	390
Krasnodar	149	207
Afipsky	153	213
Tuapse	157	301
Ilsky	157	209
Slavyansk	165	155

Notes. Bearings are measured clockwise from due north, computed as the great-circle bearing from the nearest point on the January 2022 internationally recognized Ukraine–Russia border to the refinery centroid. The construction is fully observable from public data (the border geometry is fixed; the refinery coordinates are documented). The same bearing is used to project daily ERA5-Land wind components onto the approach direction (see Section 3.4). The table lists all 24 treated refineries; Krasnodar and Novo-Ufa are treated only under the broader attribution, so the baseline panel comprises the remaining 22.

The wide span of bearings across the panel is the key source of identifying variation for the daily IV. The instrument is refinery-specific because θ_i varies substantially across the panel: the same daily wind field can be aligned with movement from Ukraine toward one refinery and opposed to movement toward another. This cross-refinery heterogeneity in directional alignment, interacted with day-to-day wind variation, generates identifying variation in strike-day incidence at the daily frequency.

D Alternative approach-bearing definition

The baseline approach bearing θ_i in Section 3.4 is measured from the nearest point on the January 2022 internationally recognized Ukraine–Russia border. A natural concern is that the operationally relevant origin is the territory under Ukrainian control during

the war rather than the pre-invasion border. We assess this directly: we recompute the directional-alignment instrument $Z_{it} = \cos(\phi_{it} - \theta_i)$ from an alternative bearing, measured to the nearest point on the perimeter of Ukrainian-controlled territory as of late 2022, and re-estimate the daily IV holding the sample, fixed effects, clustering, and outcome identical to the headline specification. Only the bearing changes.

The instrument turns out to be insensitive to this choice, for a structural reason. The alignment measure captures a *direction*: across the panel, the alternative and baseline bearings differ by a median of only 4°, and their daily alignment series correlate at 0.93, even though the corresponding origin-to-target *distances* differ substantially—by up to a factor of five for the most eastern targets. Because the preferred instrument uses direction rather than distance, the border definition does not materially alter the directional-alignment measure for most refineries.

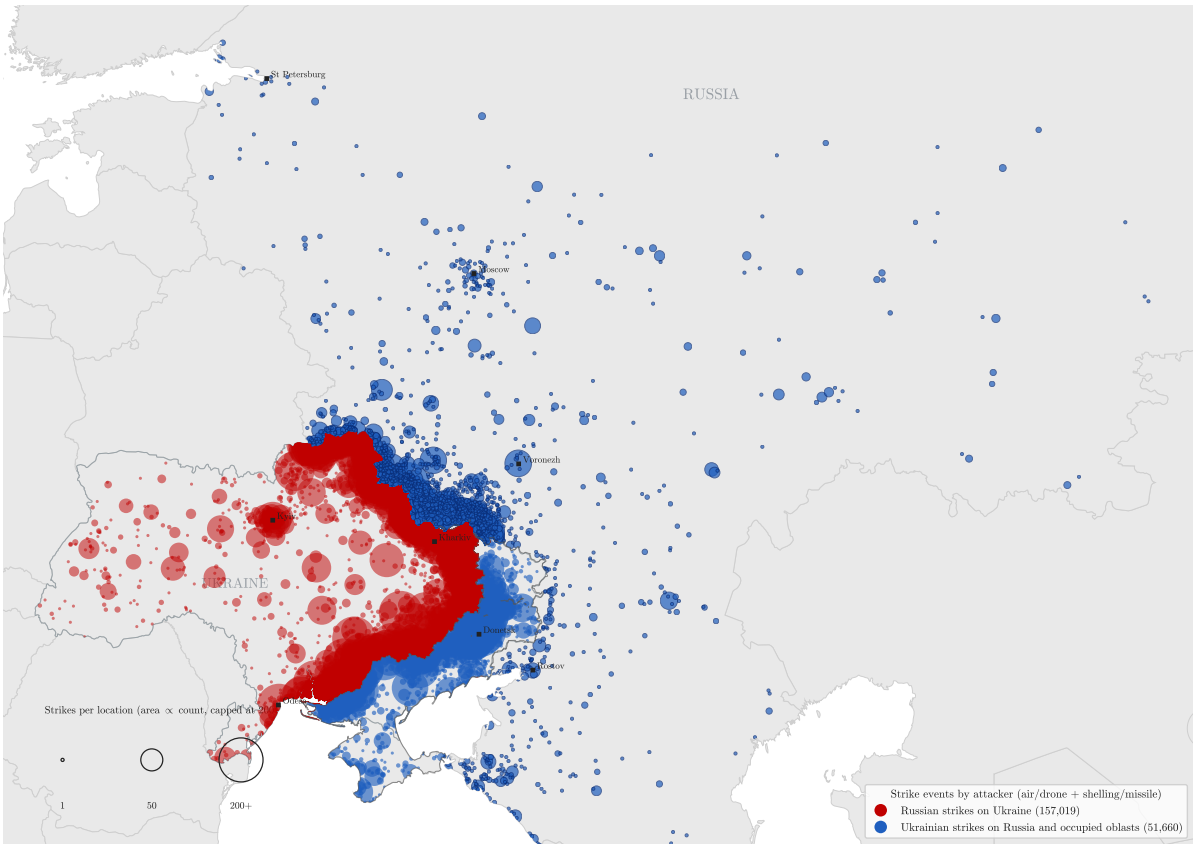
Table 10 confirms this. The reduced-form profile is the same under both bearings—essentially identical at the key horizons $h = 14$ and $h = 180$—and the first-stage F-statistics are comparable. As in the main analysis, the fire-then-damage interpretation rests on the short-run positive (fire) response at $h = 14$ and the medium-run negative (damage) response at $h = 180$, where the Anderson–Rubin intervals exclude zero under both bearings and the $h = 180$ interval is bounded in both cases. The longest horizon, $h = 270$, is reported for completeness: the baseline reduced form is smaller and statistically insignificant, and the weak first stage there makes inference less stable across bearing definitions. Overall, the daily IV results are not affected by the choice of border convention: the fire-then-damage pattern is essentially unchanged when the bearing is anchored on late-2022 Ukrainian-controlled territory instead.

Table 10: Daily IV under an alternative approach-bearing definition

Horizon (days)	RF % (base.)	RF % (alt.)	F (alt.)	AR 95% CI (alt.)
0	$+2.38^{*}$	$+3.56^{***}$	6.3	$[+5.8, +\infty)$
14	$+3.07^{***}$	$+3.08^{***}$	7.1	$[+3.5, +\infty)$
30	$+2.95$	$+2.61^{*}$	9.7	$[-1.5, +28.0]$
90	$+0.11$	$+0.77$	6.6	$[-13.2, +21.5]$
180	-4.39^{***}	-4.45^{***}	11.1	$[-26.5, -5.2]$
270	-1.81	-2.65^{**}	4.6	$[-\infty, -1.0]$

Notes. Daily IV on the baseline (CLEAN) panel, 5 km buffer, post-first-strike sample, with refinery, month-of-year, year, and day-of-week fixed effects and standard errors clustered by refinery. Each horizon h measures log radiance in a seven-day window centered h days after the strike day. The reduced-form columns report 100× the coefficient on the directional-alignment instrument under the baseline bearing—the January 2022 internationally recognized Ukraine–Russia border, as in Table 4—and under the alternative bearing. The reported first-stage F and Anderson–Rubin 95% confidence intervals are those computed from the alternative bearing; the corresponding baseline values are in Table 4. * $p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

E Bilateral strike geography



Source: authors' calculations from ACLED Partner-Tier extract; country borders from Natural Earth. A few far-east strikes fall outside the frame.

Figure 7: Bilateral strike geography: Russian strikes on Ukraine vs. Ukrainian strikes on Russia and Russian-occupied Ukrainian territory

Notes. The map plots strike events in both directions over February 2022–May 2026: Russian strikes on Ukraine (red) and Ukrainian strikes on Russia and on Russian-occupied Ukrainian territory (blue), aggregated to recorded event locations and drawn as circles whose area is proportional to the number of strikes at each location (capped at 200 so that the most intense locations remain legible). Each side's circles are clipped to the territory it targets—Russian strikes to Ukrainian-controlled territory, Ukrainian strikes to Russia and the Russian-occupied areas—so that circles do not spill across the international border or the front line, and the red/blue boundary traces the contact line. Ukrainian strikes are restricted to Russia and to the occupied oblasts of Donetsk, Luhansk, Zaporizhzhia, Kherson, and Crimea; a small number of Ukrainian strikes in government-controlled interior regions (for example around Kyiv during the 2022 invasion) are omitted. The control line is taken from DeepStateMAP; because the front moved over the period, the clipping is exact at the international border but only indicative along the internal contact line, and a few strikes in since-liberated areas are not shown. Russian strikes (157,019) outnumber Ukrainian strikes (51,660) by roughly three to one; because ACLED records discrete events rather than the quantity of ordnance delivered, this understates the asymmetry in the scale of fire. A small number of far-eastern Ukrainian strikes fall outside the frame. *Source:* authors' calculations from ACLED Partner-Tier extract; country borders from Natural Earth; control line from DeepStateMAP.

F Spatial gradient of the post-treatment effect

This appendix gives a graphical summary of the spatial gradient. Figure 8 plots the aggregate post-treatment coefficients reported in Table 1 across cumulative buffers, making the monotone decay of the decline with distance visible at a glance.

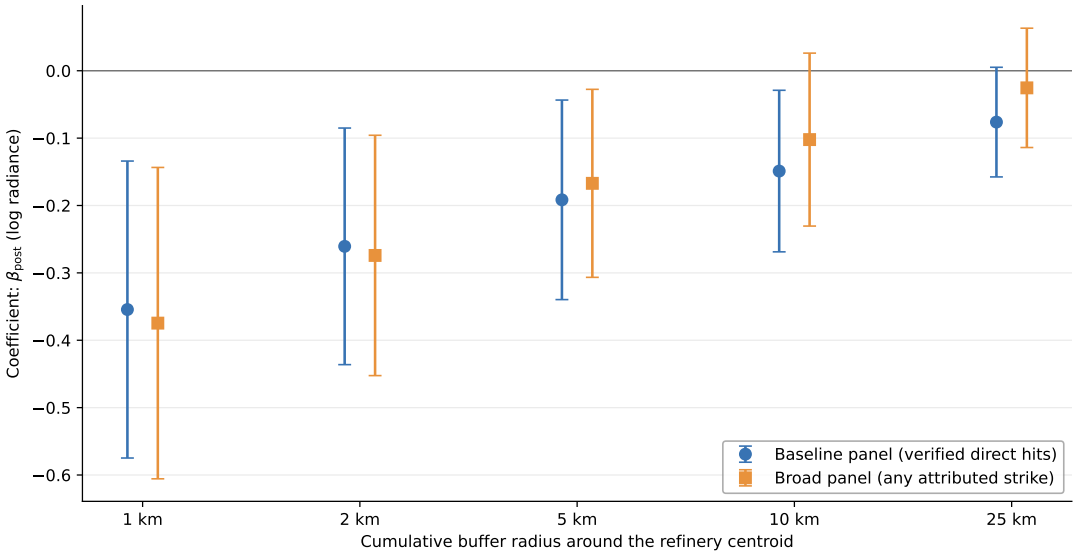


Figure 8: Spatial decay of the aggregate post-treatment effect

Notes. The figure plots the aggregate post-treatment coefficient β_{post} from equation (2) for the indicated cumulative buffer around each refinery. The dependent variable is log monthly mean quality-screened daily Black Marble radiance. All specifications include refinery and year-month fixed effects. Blue circles: baseline panel (22 treated, 7 never struck); orange squares: broad panel (24 treated, 8 never struck). Bars are 95% confidence intervals from standard errors clustered at the refinery level; underlying estimates are reported in Table 1. *Source:* authors' calculations.

G Sensitivity diagnostics and auxiliary IV estimates

This appendix collects the sensitivity analyses and auxiliary estimates referenced in Sections 4 and 5: the Rambachan–Roth bounds on violations of parallel trends, the leave-one-out diagnostics, the measurement and functional-form checks, the control-set robustness, the temporal placebo, and, for the daily IV, the binned-outcome-window estimates and the 2SLS point estimates corresponding to the Anderson–Rubin intervals in the body.

G.1 Sensitivity to deviations from parallel trends

Table 11 reports Rambachan–Roth relative-magnitudes bounds for the aggregate post-treatment effect. The honest confidence set allows post-treatment violations of parallel trends up to $\bar{M}$ times the pre-period deviations. The breakdown value is the threshold at which zero first enters the honest confidence set.

Table 11: Rambachan–Roth sensitivity of the aggregate effect (5 km buffer)

	β_{post}	(SE)	Honest CI, $\bar{M} = 0.5$	Honest CI, $\bar{M} = 1$	Breakdown $\bar{M}$ (max)	(median)
Baseline panel	−0.192	(0.076)	[−0.372, −0.011]	[−0.404, +0.021]	0.68	1.31
Broad panel	−0.167	(0.071)	[−0.355, +0.020]	[−0.403, +0.068]	0.29	0.53

Notes. Relative-magnitudes sensitivity in the spirit of Rambachan and Roth (2023), applied to the aggregate post-treatment coefficient of equation (2). The benchmark deviation is the maximum (respectively, median) absolute pre-period event-study coefficient over $k \in [-6, -2]$; in the baseline panel these are 0.064 and 0.033 log points. The honest confidence interval widens the conventional interval by $\bar{M}$ times the benchmark deviation. The breakdown value is the value of $\bar{M}$ at which the honest interval first includes zero. In the baseline panel, the aggregate effect remains distinguishable from zero at $\bar{M} = 0.5$.

G.2 Influential units

Table 12 summarises leave-one-out estimates of the aggregate effect in the baseline panel: dropping one treated refinery at a time (22 re-estimations) and one treatment cohort at a time (11 re-estimations).

Table 12: Leave-one-out estimates of the aggregate effect (baseline panel, 5 km)

	Estimates	Range of β_{post}	Range of implied %	Min $\lvert z \rvert$
Full sample	1	−0.192	−17.4	2.54
Drop one treated refinery	22	[−0.243, −0.160]	[−21.6, −14.8]	2.22
Drop one treatment cohort	11	[−0.213, −0.174]	[−19.2, −16.0]	2.24

Notes. Each row summarises the distribution of the aggregate post-treatment coefficient of equation (2) across re-estimations that exclude one treated refinery (respectively, one first-strike cohort) at a time. Every re-estimation remains negative and significant at the 5% level.

G.3 Measurement and functional form

Table 13 re-estimates the aggregate effect under alternative treatments of the daily-to-monthly aggregation and of the functional form.

Table 13: Measurement and functional-form robustness (baseline panel, 5 km)

Specification	β_{post}	(SE)	Implied %	z
OLS, log radiance, unweighted (baseline)	−0.192**	(0.076)	−17.4	−2.54
OLS, log radiance, pixel-weighted monthly mean	−0.157**	(0.073)	−14.6	−2.17
WLS, log radiance, weight = valid days	−0.167**	(0.072)	−15.4	−2.34
Poisson pseudo-ML, radiance levels	−0.245**	(0.066)	−21.7	−3.72

Notes. All rows estimate the aggregate specification of equation (2) on the baseline panel at the 5 km buffer with refinery and year-month fixed effects and standard errors clustered at the refinery level. Row 2 weights daily observations by the number of valid pixels when aggregating to the month; row 3 weights monthly observations by the number of valid days; row 4 estimates a Poisson pseudo-maximum-likelihood model on mean radiance in levels. The headline range of 15–18% quoted in the text spans rows 1–3. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

G.4 Control-set robustness

The five never-struck refineries located in eastern Russia lie far from the conflict theatre; Table 14 shows that the headline estimate does not depend on their inclusion.

Table 14: Control-set robustness (baseline panel, 5 km)

Control set	β_{post}	(SE)	Implied %	z	Pre-trend p-value
All seven never-struck refineries (baseline)	-0.192^{**}	(0.076)	-17.4	-2.54	0.98
Western controls only (drops 5 eastern units)	-0.182^{**}	(0.081)	-16.6	-2.24	0.98

Notes. Aggregate specification of equation (2), baseline panel, 5 km buffer. The second row excludes the five never-struck refineries in eastern Russia (Achinsk, Angarsk, Chernigov, Khabarovsk, Komsomolsk), retaining the two never-struck units in European Russia and western Siberia together with the not-yet-treated comparisons. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

G.5 Temporal placebo

Table 15 assigns each treated refinery a counterfactual first strike twelve months before the true one, discards all truly post-treatment observations, and re-estimates the aggregate specification on the pre-period.

Table 15: Temporal placebo: counterfactual first strike at $k = -12$ (5 km)

	β_{placebo}	(SE)	Implied %	z
Baseline panel	-0.055	(0.041)	-5.3	-1.33
Broad panel	-0.009	(0.051)	-0.9	-0.18

Notes. The placebo indicator equals one for treated refineries in the twelve months preceding the true first strike; all observations from the true first-strike month onward are excluded, so the regression compares the year before the strike with the earlier pre-period and the never-struck controls, with refinery and year-month fixed effects and clustering at the refinery level. Both estimates are statistically indistinguishable from zero, consistent with the flat pre-treatment event-study coefficients documented in the main text.

G.6 Daily IV: binned outcome windows

Table 16 re-estimates the daily IV reduced form with the outcome averaged over wide day-bins covering the year around the (potential) strike day, instead of the discrete horizons of Table 4, as discussed in Section 5.2. The pre-event bin is a placebo: wind on day t must not explain radiance before t.

Table 16: Daily IV: reduced form with binned outcome windows, 5 km buffer

Outcome window (days)	RF (%)	(SE)	p	N
Panel A: baseline panel				
Placebo $[-45, -10]$	+0.07	(0.65)	0.919	11,246
$[0, 14]$	+1.35	(1.17)	0.250	11,050
$(15, 90]$	+4.10***	(1.32)	0.002	9,880
$(91, 180]$	−2.05**	(0.88)	0.020	8,527
$(181, 270]$	−2.43**	(1.04)	0.020	7,064
Panel B: broad panel				
Placebo $[-45, -10]$	−0.19	(0.60)	0.751	12,412
$[0, 14]$	+1.23	(1.15)	0.285	12,172
$(15, 90]$	+4.30***	(1.16)	< 0.001	10,893
$(91, 180]$	−2.16**	(0.87)	0.013	9,454
$(181, 270]$	−2.40**	(0.97)	0.013	7,857

Notes. Reduced-form coefficient (in percent) of the binned outcome on directional wind alignment $\cos \delta_{it}$, on the post-first-strike estimation sample at the 5 km buffer. The outcome in each row is the mean of daily log radiance over the indicated day window relative to day t, requiring at least one quarter of the window's days to have valid observations. Refinery, month-of-year, year, and day-of-week fixed effects; standard errors clustered at the refinery level in parentheses. The pre-event window $[-45, -10]$ is a placebo: it ends before the seven-day outcome windows of Table 4 begin, so no mechanical overlap with day t remains. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

G.7 Daily IV: 2SLS point estimates

Table 17 reports the conventional 2SLS point estimates and standard errors corresponding to the reduced forms and Anderson–Rubin intervals of Table 4. As discussed in Section 5.2, these coefficients rescale the reduced forms by the wind-sensitive margin of strike incidence and should not be interpreted as average effects of typical strikes; the reduced form and the Anderson–Rubin intervals remain the quantities of interest.

Table 17: Daily IV: 2SLS point estimates, 5 km buffer

Horizon h (days)	Baseline panel 2SLS	(SE)	Broad panel 2SLS	(SE)
0	+10.56*	(6.19)	+9.55*	(5.51)
14	+10.60**	(3.38)	+9.94**	(3.34)
30	+9.77	(6.08)	+9.65*	(5.30)
90	+0.44	(6.99)	−1.24	(7.61)
180	−10.31**	(3.25)	−8.47**	(2.83)
270	−5.94	(4.13)	−6.45*	(3.34)

Notes. 2SLS coefficient on the strike-day indicator instrumented by directional wind alignment Z_{it}, in log-radiance units, as in the AR intervals of Table 4; same sample, outcome construction, and fixed effects. Conventional standard errors clustered at the refinery level in parentheses, with stars denoting conventional (Wald) significance. These conventional significance levels should be read cautiously where the first stage is weak; the horizon-specific first-stage $F(h)$ is reported in Table 4, and the weak-instrument-robust Anderson–Rubin intervals are the preferred inference. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

H Heterogeneity by target type at the 1 km buffer

Section 5.3 reports the cross-target comparison at the 5 km buffer. Table 18 repeats the exercise at the 1 km buffer—the innermost measured ring, where the refinery effect is sharpest—with the same satellite product, quality screening, and specification for every category.

The contrast sharpens. The refinery contraction rises to −29.8% (significant at 1%), while no other category develops a significant decline. Two cells deserve comment. The military estimate, −7.7%, falls just short of significance ($t = 1.63$): a small local dimming around repeatedly struck bases is plausible but not established, and even at face value it is about a quarter of the refinery effect. Power infrastructure fails the pre-trend diagnostic at this buffer ($p = 0.02$), so its positive, insignificant estimate is not interpreted.

Table 18: Target heterogeneity at the 1 km buffer

Target category	β_{post}	(SE)	Implied %	Sites	Pre-trend p-value
Oil refineries (baseline panel)	-0.354***	(0.112)	−29.8	29	0.63
Military targets	-0.080	(0.049)	−7.7	46	0.22
Fuel depots	+0.037	(0.095)	+3.7	10	0.42
Other industrial plants	-0.027	(0.042)	−2.6	22	0.93
Power infrastructure	+0.164	(0.156)	+17.8	17	0.02
Transport nodes	+0.008	(0.093)	+0.8	11	0.24

Notes. Aggregate post-treatment coefficient from equation (2) estimated separately by target category, as in Table 6, but at the 1 km cumulative buffer (the $(0.5, 1]$ km ring: every buffer excludes the inner 500 m disk). Log monthly mean of quality-screened daily Black Marble radiance, unit and year-month fixed effects, standard errors clustered at the unit level in parentheses. The pre-trend column reports the p-value of the joint test that the event-time coefficients $k \in [-6, -2]$ are zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I Weather-feasibility audit of the daily IV first stage

Section 4.2 states that directional wind alignment $Z_{it} = \cos(\phi_{it} - \theta_i)$ is the preferred daily instrument among the weather-based candidates we evaluate, and that richer specifications adding further weather variables do not improve the first stage. Table 19 reports the underlying audit on the post-first-strike estimation sample of the baseline panel. In each specification the strike-day indicator is regressed on Z_{it} together with the listed weather variables, entered additively as separate instruments—not interacted with Z_{it}; the table reports the coefficient on Z_{it} and the joint first-stage F-statistic of the full instrument set.

Table 19: First-stage audit of weather-feasibility instruments

Instrument set	Alignment Z_{it}	Joint first-stage F
S1: directional alignment only (baseline)	+0.0027***	12.0
S2: + heavy precipitation	+0.0027***	9.0
S3: + crosswind and maximum wind	+0.0032***	4.1
S4: + snowfall	+0.0028***	7.1
S5: + freezing at target	+0.0027***	6.0
S6: + all corridor weather	+0.0030***	3.4
S7: + target-cell weather	+0.0031***	4.4

Notes. Each row regresses the binary strike-day indicator on the indicated instrument set, on the post-first-strike estimation sample of the baseline panel, with refinery, month-of-year, year, and day-of-week fixed effects and standard errors clustered at the refinery level. The coefficient on directional alignment Z_{it} is reported; the joint first-stage F is the Wald statistic for the full instrument set. Snowfall is the only added variable individually significant when entered on its own (S4): a one-standard-deviation increase in snowfall is associated with about a 0.1 percentage-point reduction in daily strike probability ($p < 0.05$), consistent with reduced flight feasibility, but the effect is not robust once other weather variables are included, and it raises an additional exclusion concern because snow cover affects satellite scene reflectance directly—the class of weather-instrument violations discussed by Mellon (2025). Directional alignment alone has the strongest joint first stage; every weather addition reduces it. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Directional alignment also captures direction rather than wind magnitude. As a robustness check, Table 20 replaces it with a magnitude-weighted tailwind, wind speed$_{it} \times Z_{it}$. The reduced-form profile is the same fire-then-damage sequence—a positive response at short horizons and a negative response at six months—in both panels, so the result does not depend on the directional-only construction; the directional measure is preferred because it is less tied to broad wind systems that may affect satellite measurement.

Table 20: Robustness: magnitude-weighted tailwind, reduced form (5 km buffer)

Instrument	First-stage F	RF $h = 14$ (%)	RF $h = 180$ (%)
Panel A: baseline panel			
Directional alignment Z_{it}	12.0	+3.07***	-4.39***
Magnitude-weighted tailwind	8.6	+0.94***	-1.56***
Panel B: broad panel			
Directional alignment Z_{it}	10.1	+2.86***	-4.03***
Magnitude-weighted tailwind	10.8	+0.88**	-1.43***

Notes. Reduced-form coefficient (in percent of radiance) of the horizon-h outcome on the instrument, post-first-strike estimation sample, 5 km buffer, with refinery, month-of-year, year, and day-of-week fixed effects and refinery-clustered standard errors. The magnitude-weighted tailwind multiplies directional alignment by wind speed (m s^{-1}); its coefficients are therefore on a per-unit-speed scale and not directly comparable in level to the directional measure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Repeated-strike intensity diagnostic

A natural question is whether the headline monthly contraction is driven by refineries struck many times rather than by entry into the campaign as such. We address this descriptively. Repeated-strike intensity is endogenous and post-treatment: Ukraine re-strikes refineries on the basis of strategic value, reachability, and the outcome of earlier strikes, and a refinery's cumulative number of verified direct hits mechanically increases with time since it entered the campaign (in the post-treatment sample its correlation with event time is about 0.45). The exercise below is therefore descriptive and is not a causal dose-response estimate of the marginal effect of an additional strike.

Table 21 reports three monthly DiD specifications at the 5 km buffer. Column (1) reproduces the headline aggregate as a sanity check. Column (2) restricts the treated group to the six refineries that receive a single verified direct hit over the sample period, compared with the seven never-struck controls. Column (3) is a within-treated, post-entry regression on treated refinery-months, adding event-time-bin indicators (months $[0, 3]$, $[4, 6]$, $[7, 12]$, $[13, +)$) to refinery and year-month fixed effects, and reporting the coefficient on the lagged cumulative count of verified direct hits.

Table 21: Repeated-strike intensity diagnostic (monthly DiD, 5 km buffer)

	(1) Baseline	(2) Once-struck	(3) Within-treated
β_{post}	-0.192^{**}	-0.190^{***}	
	(0.076)	(0.066)	
Lagged cumulative direct hits			+0.030
			(0.021)
Implied % change	−17.4	−17.3	—
Refinery FE	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes
Event-time-bin controls	No	No	Yes
Refineries	29	13	22
Observations (refinery-months)	1,279	553	409

Notes. Monthly DiD at the 5 km cumulative buffer; dependent variable is log monthly mean quality-screened Black Marble radiance. Standard errors clustered at the refinery level in parentheses. Column (1) is the baseline aggregate of equation (2) (22 treated + 7 never-struck controls). Column (2) keeps only the six treated refineries with exactly one verified direct hit, against the seven never-struck controls. Column (3) uses treated refinery-months in the post-entry regime ($k \geq 0$), with refinery and year-month fixed effects and indicators for event-time bins $[0, 3]$, $[4, 6]$, $[7, 12]$, $[13, +)$; the reported coefficient is on the number of verified direct hits accumulated through the previous month. Cumulative strike intensity is endogenous and post-treatment and is strongly correlated with time since entry (about 0.45 in this sample); the table is descriptive and is not a causal dose-response estimate. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Two patterns emerge, both reinforcing the campaign-entry interpretation of the headline estimate. First, the contraction is already present among refineries that receive a single verified direct hit: the six once-struck refineries show a decline of −17.3% (column 2), essentially identical to the headline −17.4% estimate and slightly more precisely estimated, so the result is not driven by the heavily struck refineries. Second, conditional on entry, cumulative repeated strikes add no stable within-treated gradient: once refinery, year-month, and event-time-bin fixed effects absorb the timing of exposure, the coefficient on the lagged cumulative direct-hit count is small, statistically insignificant, and if anything positive (column 3). This pattern is consistent with interpreting the monthly estimand as the cumulative local effect of entering the strike campaign, rather than as a function of the number of strikes received.