HEC Paris
Master Thesis International Finance

Multi-Modal Satellite Monitoring of Open-Pit Mine Productivity: Synergizing Atmospheric NO_2, Aerosols, and Radar-Based Volumetric Sensing

Yannish Sagoenie*
Ruben de Laat

Academic advisor:	Prof. Thierry Foucault
Program:	International Finance
Date of submission:	16th June 2026

Keywords: Sentinel-5P, satellite data, open-pit mines, alternative data

*Y.K. Sagoenie
HEC Paris, 1 Rue de la Libération,
78350 Jouy-en-Josas, France
E-mail: yannish.sagoenie@hec.edu

R.R.M. de Laat
HEC Paris, 1 Rue de la Libération,
78350 Jouy-en-Josas, France
E-mail: ruben.de-laat@hec.edu

Acknowledgement

We would like to express our deepest gratitude to our academic supervisor, Professor Thierry Foucault, for his invaluable guidance, patience, and insightful feedback throughout the development of this thesis. His expertise in financial economics and his encouragement to explore novel frontiers in alternative data were fundamental to the completion of this research. A special thanks is extended to the European Space Agency (ESA) for providing the financial and structural support that made this satellite-based research possible through the Network of Resources (NoR) framework. We are also highly grateful to the EarthConsole platform and its technical team. Their continuous support, prompt assistance, and expertise were indispensable in overcoming the computational bottlenecks, algorithm complexities, and data processing issues encountered during the interferometric analysis.

Abstract

This paper tests whether satellite-derived signals from the Sentinel constellation contain real-time information about open-pit mining firm earnings ahead of quarterly disclosures. Using a hand-collected panel of 39 mines across 17 corporations spanning Q2 2018 to Q4 2025, we estimate two-way fixed effects and first-differences models relating tropospheric NO_2 and aerosol index signals from Sentinel-5P, net of spatially matched control areas, to quarterly reported production and firm-level earnings surprises. The headline first-differences estimate for commodity production is positive and significant, robust across ten alternative specifications, and strongest for iron ore and high-retention arid environment mines, establishing the diesel combustion channel as the operative physical mechanism linking satellite observations to extraction activity. The aerosol index produces a null result across all specifications. At the firm level, above-median NO_2 signals are associated with more than twice the odds of beating the analyst consensus EPS estimate for high-coverage firms, and an encompassing regression confirms that the satellite-derived production forecast contains predictive content incremental to the Bloomberg analyst consensus. A long-short trading strategy based on this signal yields announcement-day returns indistinguishable from zero, consistent with satellite information being at least partially priced in before the earnings announcement, while a pre-announcement contrarian strategy generates positive risk-adjusted returns, consistent with sophisticated investors overreacting to a freely discoverable signal. As a supplementary cross-check, we construct pit-volume estimates from Sentinel-1 radar interferometry on a subsample of mines, yielding directionally consistent results. These findings establish tropospheric NO_2 as a statistically detectable and practically deployable alternative data signal for trading the mining sector, with informational value concentrated at smaller high-coverage firms operating in low-cloud arid environments.

Contents

1 Introduction

Investors in mining equities face a fundamental information problem. Production is the primary driver of cash flows for resource firms, yet it is disclosed on a quarterly basis with a material lag between physical extraction and corporate reporting. Conventional sources of forward-looking information, including analyst forecasts, management guidance, and industry surveys, rely on the same delayed disclosures. Standard alternative data cannot bridge this gap: parking-lot imagery is irrelevant for industrial extraction, and credit-card transactions carry no information about commodity output. A commodity-focused analyst covering mining equities currently waits four to eight weeks after quarter-end for production disclosures before constructing earnings forecasts; while sell-side consensus estimates exist, they are anchored to the same delayed operational data, leaving no conventional basis for directional positioning that is independent of the corporate reporting cycle. This thesis asks whether freely available satellite-derived atmospheric signals can close that gap, providing investors with real-time information about mining production and, by extension, about likely earnings surprises before they are publicly reported.

The European Space Agency's Sentinel-5P satellite, operational since 2018, offers a physically grounded solution. Mining operations generate measurable atmospheric activity: diesel-powered haul trucks, drills, and excavators emit nitrogen dioxide (NO_2), while blasting and overburden displacement mobilise airborne aerosols. Because tropospheric NO_2 persists for only several hours before chemical degradation, concentrations detected above a mine reflect its current operational tempo rather than accumulated regional emissions. Sentinel-5P measures these signals daily at 3.5×5.5 km resolution, sufficient to isolate individual mine footprints, and makes the resulting data openly accessible. If these atmospheric signals correlate systematically with operational intensity, they offer a real-time, cost-free window into mining production that quarterly disclosures cannot provide.

This thesis assesses whether Sentinel-5P-derived atmospheric signals contain informational content about mining production and whether that physical relationship transmits into financially exploitable information at the firm level. Using a hand-collected panel of 39 mines from 17 major mining companies, covering quarterly observations from 2018 through 2025, we test whether tropospheric NO_2 and aerosol index measurements over each mine, net of regional atmospheric variation captured by spatially matched control areas, predict reported production after netting out mine-level and time-level fixed effects. The sample includes copper, gold, and iron ore operations across 17 countries. We then assess whether the satellite-production relationship translates into predictability of firm-level earnings surprises by comparing satellite-derived production estimates with quarterly analyst consensus forecasts obtained from Bloomberg, and evaluate whether any predictive content is exploitable in a systematic trading strategy around earnings announcements. An exploratory supplementary analysis using volumetric measurements derived from Sentinel-1 InSAR on a subsample of mines serves as a physically independent cross-check.

We find that tropospheric NO_2 contains statistically detectable information about within-mine production variation under the preferred first-differences specification, with the headline estimate significant at the one percent level ($\hat{\beta} = +2{,}690$, $p = 0.009$, $N = 701$), robust across ten alternative specifications and strongest for smaller, iron ore mines and high-retention arid-environment operations. The aerosol index produces a clean null result throughout, indicating that the diesel combustion channel rather than the dust channel is the operative mechanism. At the firm level, above-median NO_2 signals are associated with more than twice the odds of beating the analyst consensus EPS estimate for high-coverage firms (odds ratio 2.55, $p = 0.018$). An encompassing regression confirms that the satellite-derived production forecast contains predictive content incremental to the Bloomberg consensus, with the forecast coefficient significant at the five percent level after controlling for the consensus estimate, indicating that the signal carries information about realised earnings that sell-side analysts have not yet incorporated. A long-short trading strategy based on this signal yields announcement-day returns indistinguishable from zero, consistent with the satellite information being at least partially priced into equities prior to the announcement. Because this freely discoverable signal appears to be overweighted by sophisticated investors ahead of announcements, a pre-announcement contrarian strategy that fades the signal direction generates positive mining-adjusted returns of $+1.72$ percent per trade ($p = 0.039$) over the $[-1, +1]$ window, consistent with this overreaction creating a reversal premium. The InSAR volume analysis on a subsample of mines yields directionally consistent results that hint at the potential of radar-based volumetric sensing as a physically independent cross-check on extraction activity; the commodity production specification reaches conventional significance under bootstrap inference while the material moved specifications fall short, and both point to a genuine physical mechanism whose full detection awaits a larger dataset of coherent displacement estimates.

A substantial literature documents the predictive power of alternative data for firm-level earnings, with applications including consumer reviews (Green et al., 2019), retail satellite imagery (Katona et al., 2025), and online attention measures (Da et al., 2011). However, applications to industrial production and commodity-sector earnings remain largely unexplored, and the question of how a freely available atmospheric signal propagates through the price discovery process has not been examined in this setting. We focus on four contributions. First, we establish that above-median NO_2 signals predict earnings beats with an odds ratio of 2.55 and that the satellite forecast contains information incremental to the analyst consensus, extending the alternative data earnings-predictability literature from consumer-facing sectors to industrial commodity extraction. Second, we provide evidence on how this freely available signal is absorbed into equity prices: the announcement-day null combined with positive pre-announcement contrarian returns is consistent with sophisticated investors partially incorporating and subsequently overweighting the signal, generating a reversal premium that a contrarian strategy captures. Third, we identify the mine and commodity types for which the signal is most informative, providing practical guidance for practitioners: the signal is most reliably deployed at high-coverage firms operating iron ore or copper mines in low-cloud arid environments. Fourth, we document the methodological choices specific to Sentinel-5P industrial monitoring, including the appropriate spatial aggregation relative to mine footprints and the control-area subtraction procedure that isolates mine-specific atmospheric signatures from regional variation.

These findings have direct implications for investment practice. For a long-short equity strategy, the satellite framework provides directional positioning ahead of quarterly disclosures; for a commodity desk, it provides an independent real-time read on supply-side fundamentals;

for a corporate-credit analyst, it provides early signals of operational disruption. The signals are most reliably deployed for the mine and commodity types this thesis identifies as having the strongest signal-production link, and the pre-announcement window rather than the announcement date itself is where the trading opportunity is concentrated.

The remainder of this thesis is organized as follows. Section 2 reviews the related literature on alternative data in finance and satellite imagery as an information source for capital markets. Section 3 describes the construction of the production, satellite and financial datasets. Section 4 presents the empirical methodology, including the construction of net atmospheric signals and the panel specifications. Section 5 reports results on the production-prediction analysis, extends the analysis to firm-level earnings predictability, and discusses the results of implementing a trading strategy. Section 6 concludes with a discussion of limitations and directions for future work.

2 Literature Review

This section reviews the literature underpinning the empirical framework and signal choices of this study. Section 2.1 establishes the theoretical and empirical foundations of alternative data in finance. Section 2.2 narrows to satellite observation as a distinct class of alternative data, motivating the departure from visible-spectrum imagery toward physical process signals. Section 2.3 documents the diesel combustion mechanism underlying the NO_2 signal and its empirical validation against reported mine output. Section 2.4 examines aerosol concentration as a complementary dust-based signal and its meteorological complications. Section 2.5 establishes how repeat-pass InSAR recovers surface displacement as a geometrically direct record of material excavation.

2.1 Alternative data in finance

A growing body of research investigates the use of alternative data sources in predicting various outcomes such as returns or earnings. The theoretical motivation of utilizing alternative data for prediction stems back to Grossman and Stiglitz (1980), whose foundational paradox established that fully informationally efficient prices are impossible when information is costly: in equilibrium, traders who acquire costly signals must be compensated by trading profits, so prices reflect private information only "partially". Alternative data is precisely the modern instantiation of this costly information. A growing theoretical literature studies how the rise of alternative data impacts this equilibrium. Dugast and Foucault (2018) model the relation between low precision signals (e.g., raw data) and subsequent precise signals (e.g., fundamental analysis), and conclude that the decrease in costs of producing low precision signals raise short-run informativeness but reduce long run price informativeness. In their sequel, Dugast and Foucault (2024) formalize the competition between quantitative data miners and fundamental asset managers, in which it is shown that data abundance raise price informativeness but widen the performance dispersion between across asset managers. Moreover, data abundance contributes to disproportionate growth in equity capital markets (Begenau et al., 2018). Specifically, Begenau et al. (2018) show that larger firms generate more analyzable data, and the availability of alternative data lowers their cost of capital, thereby amplifying their relative growth advantage.

Empirically, alternative data has now shown to predict fundamentals across a wide set of domains. Froot et al. (2017) use proprietary data of consumer activity and demonstrate that real-time transaction proxies can forecast quarterly sales, yielding average excess returns of 3.4% over the earnings announcement window. H. Chen et al. (2014) show that user-generated opinions from Seeking Alpha predict earnings beyond what financial analysts capture. In line with H. Chen et al. (2014), Bartov et al. (2017) conclude that pre-announcement Twitter content contains analogous predictive power. Green et al. (2019) establish that changes in Glassdoor employer ratings forecast one-quarter-ahead earnings surprises and generate roughly 3% annualised alpha. Investor-attention measures based on Google search behaviour (Da et al., 2011) and textual sentiment in financial media (Tetlock, 2007; Tetlock et al., 2008) similarly contain information about cashflows and returns.

Another important research field is the question how alternative data shifts the production of financial information. Zhu (2019) uses the staggered availability of consumer-transaction and satellite imagery datasets as a quasi-experiment, finding that alternative data increases price informativeness and disciplines managerial rent extraction. Grennan and Michaely (2020) show that FinTech platforms aggregating alternative data partly substitute for sell-side analyst research, while Cao et al. (2024) demonstrate that AI-augmented analysts armed with alternative data outperform purely human peers. Bai et al. (2016) provide the broader backdrop: price informativeness at three-to-five-year horizons has risen secularly since 1960, concentrated in firms with greater institutional ownership.

A central caveat is that these informational benefits are sharply horizon-dependent. Dessaint et al. (2024) formalise this insight: because forecasting short- and long-term cash flows compete for limited analyst attention, the arrival of cheap, short-horizon alternative data optimally reallocates effort toward the near term. The result is a "horizon effect" (empirically validated using analyst exposure to social-media data) in which short-term forecast informativeness rises while long-term informativeness falls. This finding reframes the policy and welfare debate. Alternative data does not uniformly enhance price discovery; it redistributes informational precision across horizons and across investors with heterogeneous data access (Katona et al., 2025).

Two gaps emerge from this body of work. The empirical literature has concentrated on consumer-facing sectors such as retail, restaurants, airlines, and consumer technology, where transaction, search, and review data are abundant. Industrial sectors remain comparatively underexplored, despite exhibiting precisely the conditions under which alternative data should be most valuable: opaque and geographically dispersed production, quarterly rather than continuous disclosure, and thin analyst coverage. Mining is a particularly stark case. The present study applies this same earnings-announcement logic to the mining sector, asking whether satellite-derived production signals observable prior to quarterly disclosures predict EPS surprises and generate abnormal returns around announcement dates. A second gap concerns the dominant non-consumer modality. Satellite observation in finance has so far been almost synonymous with visible-spectrum imagery of parking lots, nighttime lights, and cargo movements, leaving the informational content of atmospheric and geophysical measurements largely unexamined.

2.2 Satellite data as a form of alternative data

Imagery captured by earth-observation satellites as an alternative data input for economic and financial inference remains an important feature for academics and practitioners. The foundational contribution is Henderson et al. (2012), who use nighttime light intensity as a proxy for subnational economic activity, establishing that what is visible from orbit can serve as a credible and politically unmediated substitute for official statistics in regions where conventional data are unreliable or strategically manipulated. The approach was validated contemporaneously by X. Chen and Nordhaus (2011) and has since been extended across development economics, spatial analysis, and financial research (Donaldson & Storeygard, 2016), situating satellite imagery within the broader turn toward big-data-driven empirical work (Einav & Levin, 2014; Varian, 2014). Within finance specifically, Katona et al. (2025) demonstrate that parking-lot occupancy derived from commercial satellites predicts retailers' quarterly sales and generates abnormal returns around earnings announcements, with gains accruing disproportionately to sophisticated investors. Zhu (2019) shows that access to such signals raises price informativeness and disciplines managerial opportunism, while Mukherjee et al. (2021) exploit cloud-cover variation over U.S. crude-storage hubs and Chinese manufacturing centres to show that satellite-derived estimates substitute for government macro releases, attenuating the price impact of official announcements (see also Hao et al., 2024). Beyond imagery, AIS vessel-tracking data offer a complementary window onto global trade flows (Brancaccio et al., 2020), and machine-learning methods applied to daytime imagery now recover local welfare levels (Jean et al., 2016) and field-level crop yields (Burke & Lobell, 2017).The signal that has attracted the most attention, nighttime luminosity, is nonetheless subject to well-documented measurement constraints: sensor saturation in dense industrial and urban environments causes the brightness of precisely the most economically active areas to be systematically truncated (Bluhm & Krause, 2022), while successive satellite generations suffer from calibration drift that complicates long-run time-series comparisons (Gibson et al., 2021). Crucially for extractive industries, luminosity is inherently insensitive to daytime industrial activity, which constitutes the majority of open-pit mining output, and provides no information about the physical processes, material volumes, or atmospheric byproducts that characterise mine operations. These limitations motivate the development of alternative physical signals that are measured during working hours, scale directly with operational throughput, and remain within the dynamic range of current spaceborne instrumentation. The present study applies three such signals to the monitoring of open-pit copper, gold, and iron ore mining: tropospheric nitrogen dioxide, which is emitted by the diesel combustion of the haulage fleet and scales with total material moved; particulate dust loading, which is mobilised by blasting and vehicle traffic across exposed bench surfaces; and surface displacement volume, which is recoverable from radar interferometry and provides a geometrically direct record of excavation. Together these constitute a multi-signal satellite framework for inferring mine production in precisely the remote and institutionally opaque environments where conventional data are most deficient.

2.3 Nitrogen dioxide as a proxy for production intensity

Nitrogen dioxide (NO_2) arises in open-pit mining from two concurrent physical processes: the high-temperature combustion of diesel fuel by haulage trucks, excavators, and loaders; and the detonation of commercial explosives during blasting cycles. Both inject NO_X directly into the near-surface boundary layer in quantities that scale with the volume of material handled and the total energy consumed. Critically, NO_2 persists in the troposphere for only several hours before reacting with the hydroxyl radical, ensuring that detected column enhancements remain spatially proximal to their source (Martinez-Alonso et al., 2023). This brevity makes NO_2 peculiarly well suited to industrial source attribution: concentrations measured overhead a mine reflect its current operational tempo rather than the accumulated regional combustion legacy, an advantage unavailable for longer-lived tracers such as CO_2. The link between NO_2 and economic output has been formalised beyond the mining context by Parubets and Naito (2025), who demonstrate using prefecture-level Japanese GDP data that a one percent increase in tropospheric NO_2 is associated with a 0.614 percent increase in mining gross domestic product, the strongest single-sector elasticity in their analysis, reflecting the energy intensity that distinguishes extractive operations from other industrial activities.

The technical instrument enabling facility-scale monitoring is TROPOMI aboard ESA's Sentinel-5P, which delivers tropospheric vertical column densities at 3.5×5.5 km ground resolution with daily near-global revisit, substantially surpassing the spatial fidelity of predecessors such as OMI and SCIAMACHY. This resolution first permitted attribution of NO_2 enhancements to individual point sources rather than broad regional emission regions (Beirle et al., 2021). A refined global catalogue by Beirle et al. (2023) subsequently expanded confirmed source detections to 1,139 facilities, with point-source correlation coefficients of 0.81 and 0.64 validated against governmental emission registers for Germany and the United States respectively. The present study adopts a background-subtraction approach, differencing mean NO_2 columns over the mine footprint against a spatially matched upwind control area, a procedure that is transparent, computationally tractable, and well suited to detecting relative change over time rather than absolute emission rates.

The empirical basis for using TROPOMI-derived NO_2 as a mine-production proxy was established by Martinez-Alonso et al. (2023), who demonstrated that annual NO_X emission estimates at four large open-pit copper and copper-cobalt operations in the Central African Copperbelt correlate strongly with independently reported production, with Pearson coefficients ranging from 0.73 to 0.83. Crucially, the best-performing production proxy was not refined metal output, which is sensitive to ore grade and processing efficiency, but total ore and waste mined: a volume-based metric mechanically proportional to haulage energy and therefore to NO_X emissions. This implies that the satellite signal tracks physical material throughput rather than metallurgical yield, and that its predictive relationship with production is strongest where ore-grade variability is limited. A salient caveat is that this relationship is site-specific and can drift as equipment efficiency evolves: declining NO_2 at the Lumwana mine in 2022 was traceable to a documented fleet modernisation programme rather than a production reduction, underscoring the need to recalibrate against reported production data whenever these are available (Martinez-Alonso et al., 2023).

2.4 Aerosol concentration as a proxy for production intensity

Particulate matter is generated across the entire operational cycle of an open-pit mine: drilling and perforation liberate fine rock particles at the working face; blasting injects coarse and fine fractions into the atmospheric boundary layer; shovel loading and truck tipping produce fugitive dust at material transfer points; and continuous haulage vehicle movement over unpaved roads generates road-surface dust whose magnitude scales directly with traffic intensity. Because each of these processes generates aerosol in proportion to material throughput, the column-integrated aerosol load observed above a mine is in principle an indicator of operational tempo. The machine-learning literature confirms this mechanistic link empirically: Yang et al. (2024) demonstrate at the Haerwusu open-pit coal mine that daily stripping volume and coal extraction rank among the most important predictors of measured particulate concentrations after meteorological state, and that models incorporating production intensity systematically outperform meteorology-only specifications across all particle-size fractions.

The central complication of aerosol-based production monitoring is the sensitivity of observed concentrations to meteorological state. Ma et al. (2022) document at the Anjialing coal mine that atmospheric pressure, wind speed, and humidity each exert statistically significant influences on measured particulate concentrations, with boundary-layer stability producing pronounced seasonal variation that reflects atmospheric conditions rather than changes in production intensity. Raw aerosol concentrations are therefore a joint product of emission intensity and atmospheric dispersal, and any attempt to infer the former from the latter without controlling for meteorology risks systematically biased inferences. The present study addresses this through a difference-in-differences design in which aerosol index observations over each mine footprint are differenced against a spatially proximate control area subject to the same regional meteorological conditions but not to mining emissions, so that time-varying atmospheric factors cancel in the within-period comparison, conceptually analogous to the upwind background-subtraction strategy employed for NO_2 (Martinez-Alonso et al., 2023).

The satellite instrument used is the absorbing aerosol index retrieved by TROPOMI aboard Sentinel-5P, which Darvishi Boloorani et al. (2025) employ at to produce a global atlas of dust emission sources, confirming that both natural and anthropogenic processes contribute to observed aerosol loads. For mine-level monitoring this raises an attribution challenge absent from NO_2: column-integrated aerosol products cannot spectrally discriminate between dust lofted by mining operations and natural aeolian dust mobilised from adjacent surfaces, and upwind transport can contaminate the mine-footprint signal with dust advected from considerable distances. The aerosol signal will consequently be strongest at mines situated in low and stable natural dust environments, and attenuated where regional aeolian background loads are high relative to the mining increment.

2.5 Volume displacement as a proxy for production intensity

Ground deformation constitutes a geometrically direct record of the material redistribution that open-pit mining imposes upon the landscape: as ore and overburden are excavated and deposited onto waste dumps or conveyed to processing infrastructure, the surface undergoes vertical and horizontal displacement mechanically proportional to the volume extracted. Repeat-pass

InSAR recovers this displacement signal by differencing the phase of radar backscatter measured from the same ground cell across successive satellite passes, yielding millimetre-to-centimetre precision surface velocity fields at spatial densities no ground sensor network can replicate (Escayo et al., 2022). The Small Baseline Subset (SBAS) variant achieves temporal stability by connecting image pairs with short orbital and temporal baselines, minimising phase decorrelation from rapid surface change. A practical constraint is that active extraction zones, subject to daily blasting, machinery traffic, and stockpile accumulation, frequently lose coherence within a single orbital repeat cycle and yield sparse measurement clouds, while passive elements such as waste dumps and abandoned pit walls preserve coherence over multi-year intervals and support quantitative volumetric analysis. Escayo et al. (2022) demonstrate this at the Rio Tinto complex, where Sentinel-1 C-band processing revealed four deformation anomalies, two of which on passive waste dumps were entirely unknown prior to the satellite analysis.

Converting LOS displacement rates into volumetrically meaningful estimates requires correcting for the radar illumination geometry, which is insensitive to north-south motion and conflates vertical subsidence with horizontal range-direction displacement. Where ascending and descending passes are available, full decomposition into vertical and east-west components is possible (Escayo et al., 2022), though single-geometry studies can recover meaningful displacement magnitudes through incidence-angle projection under physically motivated assumptions about the dominant deformation mode. Integrating spatially distributed displacement rates over a mine bench or waste dump face yields a volumetric deformation rate that is in principle directly comparable to surveyed material movement quantities in production accounts, providing a monitoring signal that is independent of operational reporting chains and immune to the disclosure incentives that affect voluntary and regulated production filings alike. This independence is the key financial virtue of the approach and positions InSAR-derived volume as mechanistically complementary to combustion-derived tracers such as NO_X (Martinez-Alonso et al., 2023): the two signals are generated by physically distinct processes but both scale with total material moved, so their joint behaviour provides a richer and more robust basis for inferring mine production than either signal alone.

Validation at Rio Tinto against a GNSS network operated by Atalaya Mining at the Atalaya pit's eastern wall, where block sliding produced displacements of up to eleven centimetres per year, showed high correspondence between projected LOS velocities and GNSS vectors across four co-located measurement pairs (Escayo et al., 2022). Residual mismatches reflect the inherent scale difference between point-location GNSS monuments and the forty-by-forty metre pixel averages of the radar sensor: localised high-deformation zones whose spatial extent is subpixel are smoothed into surrounding coherent areas, diluting their displacement contribution. The volumetric estimates derived here should accordingly be interpreted as spatially integrated lower bounds on true material movement rather than precise point-scale excavation rates.

3 Data

This section describes the data and selection process underlying the empirical analysis. Section 3.1 establishes the corporate and mine-level screening criteria that define the final sample of 39 mines across 17 corporations. Section 3.2 describes the production metrics manually collected from corporate disclosures, which serve as the ground-truth outcomes against which the satellite signals are evaluated. Section 3.3 details the construction and aggregation of the NO_2 and aerosol index signals retrieved from Sentinel-5P TROPOMI. Section 3.4 introduces the Bloomberg and Yahoo Finance financial data used in the earnings surprise and trading strategy analyses. Section 3.5 describes the InSAR-derived pit volume estimates constructed as a supplementary physical cross-check.

3.1 Sample selection

3.1.1 Corporate and commodity screening

We construct our sample through a multi-stage screening process designed to identify mining corporations whose operational characteristics and disclosure practices are compatible with satellite-based monitoring of atmospheric emissions. The screening proceeded in three stages: an initial commodity restriction, a filter on corporate structure and reporting transparency, and a final restriction based on the granularity of operational disclosure. We began by restricting attention to five primary global commodities: gold, silver, copper, iron ore, and coal. These sectors share two features that make them suitable for satellite-based observation. First, they involve the highest volumes of material displacement worldwide and are characterised by large-scale open-pit operations, which produce atmospheric signatures from diesel-powered haulage and overburden disturbance that are detectable from low earth orbit. Second, the open-pit geometry provides a clearly delineated industrial footprint that can be matched to the spatial resolution of Sentinel-5P retrievals (approximately 3.5×5.5 km per pixel). Underground mining operations were excluded throughout the screening on the grounds that their atmospheric emissions are vented through localised shafts rather than dispersed across the operational area, making them less reliably observable in our setting.

Within this commodity pool, we restricted the sample to publicly traded mining corporations. Public listing is essential to our methodology as it guarantees access to audited, standardised financial and operational disclosures and provides the foundation for the firm-level financial analysis we conduct as a direct extension of the production-prediction analysis. Streaming and royalty companies were excluded from the outset: such entities hold financial interests in production rather than operating mines directly, and their disclosures are oriented towards royalty receipts rather than physical extraction, rendering them inappropriate for our purposes.

A further filter narrowed the sample based on the granularity of operational reporting. We applied three principal criteria. The first concerned geographic specificity: we retained only operators publishing production figures at the individual mine-site level, eliminating those aggregating across regions or national portfolios. This restriction is essential as our research design exploits within-mine variation, which requires spatial correspondence between corporate disclosures and satellite measurements. The second related to temporal resolution: we required operators to report quarterly production for each calendar quarter independently, ensuring direct comparability with the quarterly satellite signal aggregates used in our panel. The third concerned the existence of sufficiently detailed management discussion and analysis or supplementary operational reports from which we could manually extract site-level production figures across our observation window.

Following these filters, the initial pool of roughly 70 publicly traded mining corporations meeting the commodity and listing criteria reduced to a workable set of 17 corporations with sufficient reporting granularity. The five-commodity scope reduced to three commodities in the process: too few publicly traded pure-play silver and coal operators provided mine-level disclosure at quarterly frequency, leaving an unrepresentative sample in those sectors. Conclusively, our final commodity scope consists of copper, gold, and iron ore, which together represent the three most important commodities in terms of equity market capitalisation among major listed pure-play mining operators.

3.1.2 Mine-level filtering criteria

The 17 corporations meeting the corporate-screening criteria operate approximately 209 individual mine sites across our observation window. Whereas the corporate-level screening establishes that disclosures are sufficiently detailed in principle, the mine-level filter ensures that each individual site is suitable for satellite-based monitoring in practice. We applied four filters to the 209 candidate sites.

The first filter addressed operational coherence. We excluded mines reported as a single unit by the parent corporation but physically comprising several geographically disparate pits, since the satellite signal over such a complex cannot be unambiguously attributed to a specific production figure. Put differently, the spatial-temporal correspondence between disclosed production and satellite measurement requires that each mine be a unified, geographically contiguous extraction operation. We did not have such a constraint in the corporate-screening stage as it is a property of individual assets rather than firms. The second filter excluded mines with ownership or reporting instability. Specifically, sites that switched between private and public ownership during our observation window were eliminated, as such transitions typically produce gaps or methodological inconsistencies in quarterly disclosures. In particular, joint ventures with non-disclosing private partners were excluded where the public partner's share of production could not be cleanly identified. The third filter restricted the sample to open-pit operations. As noted above, open-pit mining produces dispersed atmospheric signatures over the operational area, which are detectable by Sentinel-5P retrievals at the appropriate spatial scale. Hybrid sites combining open-pit and underground extraction were excluded on the grounds that an unobservable underground component introduces a wedge between reported production and the satellite-observable surface signal that cannot be reconciled within our framework. The fourth filter addressed conflicting land use. We excluded mines located immediately adjacent to urban centres or major industrial complexes, where regional atmospheric activity from non-mining sources would dominate the local NO_2 and aerosol signal. In particular, mines in densely populated regions where regional pollution levels would dominate mining-specific contributions were excluded; no mines in such regions survived the joint application of the operational and atmospheric-visibility filters. Furthermore, we required each mine to have a clearly dominant primary commodity, eliminating sites where production was approximately evenly split across multiple commodities and where atmospheric emission attribution to a single output would be ambiguous. Sites engaged solely in residual leaching or processing of stockpiled material, without active primary extraction, were also eliminated, as such operations do not produce the operational atmospheric signature on which our methodology depends.

3.1.3 Final sample composition

The application of the corporate and mine-level filters yields a final sample of 39 mines belonging to 17 mining corporations. The sample spans 17 countries across five continents and three commodities. Specifically, 14 of the 39 mines extract copper, 17 extract gold, and 8 extract iron ore. Geographically, the sample is concentrated in Chile (9 mines) and Australia (6 mines), reflecting these jurisdictions' dominance in global copper and iron ore production respectively, with additional substantial representation in Peru, Canada, the United States, and across West Africa.

The full sample by company, mine name, country, and primary commodity is listed in table 10 in the appendix. The 39-mine sample, whereas modest in absolute size, captures a substantial share of the operational footprint of the listed pure-play mining sector in the three commodities studied. In particular, the copper subsample includes several of the largest porphyry deposits worldwide, including Escondida (BHP), Oyu Tolgoi (Rio Tinto), and Quebrada Blanca (Teck), while the iron ore subsample includes major Pilbara assets (Yandi, Jimblebar, Robe River) operated by BHP and Rio Tinto and the Sishen-Kolomela operations of Anglo American in South Africa. The gold subsample is the most geographically diversified, with operations spanning Mauritania, Senegal, Côte d'Ivoire, Ghana, Mexico, Argentina, the Dominican Republic, Suriname, the United States, Canada, and Australia. In line with our research design, we exploit within-mine variation in the empirical analysis, such that the relatively modest cross-sectional dimension of the sample does not constrain the identification of the satellite-production relationship.

3.2 Production data

3.2.1 Data collection process and metric definitions

In this paper, we construct the production dataset through a longitudinal audit of corporate disclosures, providing the ground-truth metrics against which the satellite-derived signals are evaluated. We manually extract operational data for each of the 17 corporations identified in section 3.1, spanning the Sentinel-5P observation period from Q2 2018 to Q4 2025. The extraction process targeted three categories of corporate disclosure: quarterly management discussion and analysis reports, audited financial statements, and supplementary operational data tables. Each observation was cross-validated against these sources where available, yielding $39 \times 31 = 1{,}209$ potential mine-quarter observations.

We collected three production metrics capturing progressively narrower aspects of mining operations. Specifically, *Total Material Moved* reports the combined mass of ore and waste rock physically displaced during the quarter (tonnes); *Total Ore Mined* reports the mass of ore-grade material extracted excluding waste rock (tonnes); and *Commodity Produced* reports final processed commodity output, expressed in tonnes for copper and iron ore and in troy ounces for gold.

3.2.2 Sample coverage and descriptive statistics

Whereas *Commodity Produced* is universally disclosed across all 39 mines, significant heterogeneity exists in the reporting of upstream metrics: 14 of the 39 mines provide no *Material Moved* disclosure during the observation window, and *Ore Mined* is disclosed for only 22 of the 39 mines. In line with this disclosure pattern, we focus on *Commodity Produced* as the primary outcome variable, with *Ore Mined* examined as a secondary outcome. *Material Moved* is treated as a complementary metric reported alongside the main specifications where disclosure permits.

Table 11 reports descriptive statistics by commodity and metric. *Commodity Produced* is observed for 1,050 mine-quarters across all 39 mines, with median per-mine coverage of 31 quarters indicating effectively complete reporting for the headline metric. *Ore Mined* and *Material Moved* are observed for 522 and 597 mine-quarters respectively, with median per-mine coverage of 26 and 27 quarters among the mines that disclose them. In particular, only 2 of the 8 iron ore mines disclose upstream metrics, which restricts cross-commodity comparison for *Material Moved* and *Ore Mined.*

3.2.3 Treatment of structural breaks and zeros

Three mines exhibit documented operational features that produce structural breaks within the observation window, which we retain in the sample with their production data faithfully reported. Specifically, the Samarco iron ore operation in Brazil shows zero production for the first 10 quarters of our sample (Q2 2018 through Q3 2020) following the November 2015 collapse of the Fundão tailings dam, with production resuming in Q4 2020; the Quebrada Blanca copper operation in Chile shows three zero quarters during 2022 reflecting the transition from QB1 to QB2 phases, during which active mining was paused and operations continued solely through leaching of previously extracted stockpiles; and the Gruyere gold operation in Australia shows zero production for the first five quarters of the sample (Q2 2018 through Q2 2019) prior to its commercial production declaration in Q3 2019.

In all three cases, zeros reflect the true operational state of the mine and are not data quality issues. However, in our analysis, zero-production observations are treated as missing in the log specification, as log(0) is undefined. Furthermore, mine fixed effects absorb the time-invariant component of these mines' average production levels, partially handling the QB1-to-QB2 level shift at Quebrada Blanca, which is further explained in section 4.2.

3.3 Satellite signals

3.3.1 Data extraction and spatial geometries

We obtain daily measurements of tropospheric NO_2 and aerosols over each mine through the SentinelHub API, which provides processed access to Sentinel-5P TROPOMI Level-2 products. Specifically, we retrieve the tropospheric NO_2 column density product (in mol/m^2), which quantifies diesel combustion and high-temperature industrial activity, and the UV aerosol index (AI), which is by definition unitless. It measures absorbing aerosols including dust generated by blasting, hauling, and overburden displacement.

A central methodological challenge is the spatial standardisation of these signals to ensure that the data captured is representative of each mine's active operational footprint while minimising contamination from surrounding terrain. To address this, we apply a multi-tiered spatial strategy with four distinct geometries. Specifically, a standardised bounding box (BBox) provides the outer spatial envelope for each site; a manually delineated polygon follows the specific contours of the active pit rim; a circular buffer centred on the primary extraction zone supports radial analysis; and a separate control area is established for each mine in a position selected to share comparable geography and elevation with the target site while remaining outside the influence of the operational plume. The control area placement was verified through manual inspection in Google Earth for each of the 39 mines, with controls positioned upwind of prevailing operational wind directions where possible to minimise contamination by the mine's own emissions.

The four mine-area geometries differ primarily in spatial extent: a typical polygon delineates a few square kilometres tightly fitted to the active pit, whereas the BBox covers more area surrounding the pit due to its geometry. As we document in the section 3.3.2, this difference in spatial extent has implications for the reliability of cloud-filtered satellite retrievals.

3.3.2 Cloud filtering and the geometry selection

Sentinel-5P retrievals are sensitive to cloud cover, which biases or invalidates measurements over cloud-affected pixels. The extraction pipeline reports a cloud-impact percentage for each daily measurement, defined as the share of pixels affected by clouds within each geometry. In line with standard remote-sensing practice (see e.g., De Smedt et al. (2018)), we retain daily observations where the cloud-impact percentage does not exceed 30%.

The polygon-level cloud impact is mechanically inflated for small measurement areas: a single cloud-affected pixel represents a much larger share of a polygon's pixels than of a BBox's pixels, and since Sentinel-5P's pixel resolution (approximately 3.5 $\times$ 5.5 km) is comparable to typical mine footprint sizes, polygon-based cloud impacts become hypersensitive to any cloud presence. The empirical consequence is severe: 10 of the 39 mines retain zero usable NO_2 observations under the polygon geometry, with daily retention averaging only 42.9% across the panel. Under the BBox geometry, daily NO_2 retention rises to 59.9%, with all 39 mines retaining at least 11.9% of days. Conclusively, we adopt the BBox-Control configuration as the primary geometry for both NO_2 and AI signals, retaining the polygon geometry as a robustness specification. AI retention is substantially higher than NO_2 retention under all geometries (99.7% under BBox), which reflects the lower cloud sensitivity of AI retrievals.

3.3.3 Net signal construction and quarterly aggregation

For model fitting purposes, we construct a net signal as the daily difference between mine-area and control-area measurements, then aggregate to quarterly means. Specifically, the net signal at mine i on day t is:

$$\text{Net Signal}_{i,t} = \text{BBox}_{i,t} - \text{Control}_{i,t}$$

Here the subtraction operates daily to preserve the temporal pairing of atmospheric conditions across the two areas. The logic is that the mine-area and control-area measurements share regional atmospheric variation (weather, transport, seasonal chemistry, common shocks) but differ in their exposure to the mining operation itself; subtraction therefore isolates the mine's marginal contribution.

We validate this approach also empirically. Within-mine correlations between the BBox and control measurements are $r = +0.866$ for NO_2 and $r = +0.946$ for AI, confirming that the control area captures the bulk of regional atmospheric variation that contaminates raw mine-area measurements. Furthermore, the within-mine standard deviation of the net signal is substantially smaller than that of the raw BBox signal: subtraction reduces noise by 48% for NO_2 (from 6.73×10^{-6} to 3.48×10^{-6} mol/m^2) and by 68% for AI (from 0.262 to 0.085). In line with this evidence, net signals serve as the primary regressor in our specifications, with raw BBox measurements reported as robustness checks.

We aggregate daily net signals to quarterly means within each mine-quarter, requiring at least 30 valid (cloud-filtered) days per quarter for the aggregate to enter the panel. This threshold corresponds to approximately one-third of a calendar quarter and balances data density against sample size. Of 1,209 potential mine-quarters, 859 (71.1%) pass the threshold for NO_2 BBox and all 1,209 (100.0%) pass for AI BBox.

A residual concern is mines with persistently poor NO_2 coverage: a small number of mines retain fewer than 10 valid NO_2 quarters across the seven-year window, providing insufficient within-mine variation to support panel identification. Specifically, the mines Antamina, Highland Valley Copper, Bloom Lake, Merian, and Canadian Malartic are dropped from NO_2 regressions on these grounds, leaving 34 mines in the NO_2 analyses and 39 mines in the AI analyses. Table 12 reports descriptive statistics for the satellite signals over the resulting samples, and figure 7 visualises the data coverage by mine and quarter.

3.4 Financial data

Financial data are obtained from Bloomberg for all 17 corporations in our corporate sample. We collect, for each reporting period, the analyst consensus EPS estimate and the actual EPS reported. The consensus is sourced from Bloomberg's BEst EPS series, which represents the mean of contributing sell-side analyst estimates as of immediately prior to the earnings announcement, and thus reflects the market's pre-release expectation. We use adjusted EPS rather than reported EPS throughout, sourced from Bloomberg's BEst Adjusted EPS series, which strips out non-recurring items such as asset impairments, restructuring charges, and other one-off adjustments. For mining corporations, which frequently record large non-cash write-downs on mineral assets and project impairments unrelated to quarterly extraction activity, adjusted EPS provides a cleaner signal of the underlying operational performance that the satellite signals are designed to track. Using reported EPS would introduce noise from accounting adjustments that are orthogonal to physical production intensity, potentially attenuating the satellite-earnings relationship we seek to identify. Our primary outcome variable is the EPS surprise, defined as the percentage deviation of adjusted reported EPS from the adjusted Bloomberg consensus estimate, which renders the measure invariant to cross-currency and cross-firm scaling differences.

Daily stock return data for the six high-coverage firms are obtained from Yahoo Finance. Total returns are computed from dividend-adjusted closing prices and are used to evaluate abnormal performance around earnings announcement windows in the trading strategy analysis of section 5.7.3. The S&P 500 index and the iShares MSCI Global Metals and Mining Producers ETF serve as market and sector benchmarks respectively for the computation of abnormal returns.

Reporting frequency varies across the sample: the majority of corporations report quarterly EPS, while a subset of primarily Australian and UK-listed operators report on a half-yearly basis. For half-yearly reporters, the satellite signal is constructed as the average of the two constituent quarters' NO_2 net BBox signals and benchmarked against the Bloomberg half-year EPS surprise for the corresponding reporting period. Financial data are collected over the same window as the satellite panel, Q2 2018 through Q4 2025, with coverage effectively complete across all 17 corporations.

3.5 InSAR volume measure

As an alternative measure of physical extraction activity, we construct quarterly pit-volume estimates using InSAR displacement data obtained from the EarthConsole platform. The analysis covers six mines for which both InSAR displacement records and *Material Moved* disclosures are available throughout the observation window: Antamina, Damang, Herradura, Merian, Tarkwa, and Pueblo Viejo. While a larger set of mines in the main panel disclose *Material Moved*, InSAR-eligible sites are substantially reduced by gaps in Sentinel-1 acquisition scheduling, instrument-level data quality failures, and persistent surface coherence issues at mines located in areas of dense vegetation, high precipitation, or active surface disturbance that cause the EarthConsole SBAS algorithm to fail to converge on stable displacement estimates. For five of the six sites, displacement records span Q3 2016 through Q4 2025, providing a longer window than the main satellite panel and supporting the temporal normalisation procedure described below. Herradura is the exception, with usable InSAR records available from 2021 onwards only, owing to corrupt imagery in the earlier period.

For each site, the raw multi-temporal displacement records are spatially filtered to the operational footprint via a mine-specific geographic bounding box and temporally normalised to standard calendar quarters through linear interpolation between acquisition dates, or extrapolation using each point's long-run velocity component where target dates fall outside the coverage envelope.

Raw InSAR measurements capture displacement along the radar line-of-sight, which we convert to true vertical displacement using the directional cosine of the satellite look angle:

$$d_{\text{Vertical}} = d_{\text{Line-of-sight}} / \cos U$$

Quarterly volume changes are then computed pixel-by-pixel. Pixel dimensions of 92.39 m $\times$ 81.66 m, recovered from the geodetic spacing between measurement points converted to metres, yield a surface area of $A_{\text{pixel}} \approx 7{,}544.57\ \text{m}^2$ per pixel. The extracted volume between consecutive quarters is:

$$\Delta V = -(\Delta d_{\text{Vertical}}/100) \times A_{\text{Pixel}}$$

where the sign convention maps surface subsidence to positive extracted volume. Pixels with quarterly volume changes below 20 m^3 are set to zero to suppress atmospheric noise and baseline artefacts, and the remaining active pixels are summed to yield total net extraction volume per quarter.

4 Methodology

This section develops the empirical framework used to evaluate whether satellite-derived atmospheric signals contain information about mining production and whether that information is exploitable in financial markets. Section 4.1 introduces the empirical strategy and identification. Section 4.2 details the main specification and the first-difference robustness estimator. Section 4.3 describes the heterogeneity and dynamic-response analyses. Section 4.4 extends the framework to firm-level earnings surprises. Finally, section 4.5 describes the construction of the long-short trading strategy used to assess whether the satellite-earnings relationship is exploitable in practice.

4.1 Empirical strategy and identification

The central empirical question is whether variation in the net NO_2 and aerosol index signals over each mine area, as constructed in section 3.3, predicts variation in reported production after accounting for unobserved heterogeneity across mines and common shocks across time. The parameter of interest is the elasticity of within-mine quarterly production with respect to the quarterly net satellite signal, identified from temporal variation within each mine rather than cross-sectional variation across mines.

The panel structure of our data lends itself naturally to a two-way fixed-effects framework. Mine fixed effects absorb the time-invariant component of each mine's production scale, commodity type, unit of measurement, average grade, ore body geometry, and location-specific atmospheric background, whereas quarter fixed effects absorb common shocks that move all mines simultaneously, such as global commodity-cycle effects, broad atmospheric anomalies, and seasonal patterns shared across the panel. Identification therefore rests on the within-mine, within-quarter covariation between net satellite signals and production, after netting out everything that does not vary along these dimensions. Formally, identification of the slope parameter requires a strict exogeneity assumption conditional on the fixed effects:

$$\mathbb{E}\left[\varepsilon_{i,t} \mid \text{NetSignal}_{i,1}, \ldots, \text{NetSignal}_{i,T},\, \alpha_i,\, \gamma_t\right] = 0 \tag{1}$$

Assumption (1) requires that the residual at (i,t) is mean-independent of the entire sequence of within-mine net satellite signals after partialling out mine and quarter fixed effects. This rules out time-varying mine-specific confounders that simultaneously drive both production and net atmospheric activity, as well as feedback effects whereby past or anticipated production drives current satellite signals.

Because Sentinel-5P has covered all 39 mines uniformly since 2018, the relevant variation is temporal rather than cross-sectional and there is no introduction event from which a treatment effect could be identified. The appropriate identification strategy is therefore panel fixed effects exploiting within-mine temporal variation, rather than the difference-in-differences designs used in work studying the causal effect of data availability on capital markets (Katona et al., 2025). Our research question is informational rather than causal in that sense: we ask whether the satellite signal contains information about production and whether that information is exploitable in financial markets, not what the introduction of satellite coverage does to market behaviour.

4.2 Main specification and robustness

Our headline specification is a two-way fixed effects (TWFE) model relating quarterly production at each mine to the corresponding quarterly satellite signal:

$$\log\left(\text{Production}_{i,t}\right) = \alpha_i + \gamma_t + \beta \cdot \text{NetSignal}_{i,t} + \varepsilon_{i,t} \tag{2}$$

where $\log(\text{Production}_{i,t})$ is the natural logarithm of quarterly *Commodity Production* at mine i in quarter t; α_i denotes mine fixed effects, capturing all time-invariant mine-specific characteristics including commodity type, reporting unit, average grade, ore body geometry, and location-specific atmospheric background; γ_t denotes quarter fixed effects, absorbing all shocks that affect mines uniformly within each quarter, such as commodity-cycle effects, global macroeconomic shocks, and seasonal patterns; $\textit{NetSignal}_{i,t}$ is the quarterly mean of the daily mine-area BBox signal net of the corresponding daily control-area signal; and $\varepsilon_{i,t}$ is the residual disturbance. The parameter of interest is β, which captures the within-mine, within-quarter response of log production to the net satellite signal.

The log transformation of production serves two purposes. First, it normalises the substantial cross-commodity scale differences described in section 3.2 (copper and iron ore production in tonnes; gold in troy ounces) by converting absolute scale differences into level shifts that are absorbed by α_i. Second, it produces a coefficient that is dimensionless and interpretable as a semi-elasticity: a one-unit increase in the net satellite signal corresponds to a $100 \cdot \beta$ percentage change in production. We estimate equation (2) separately for each combination of *NetSignal* (NO_2 and AI) and dependent variable (*Commodity Production* as the primary outcome; *Ore Mined* as a secondary outcome; *Material Moved* as a complementary metric reported alongside main specifications where disclosure permits). The more limited disclosure coverage of *Material Moved*, documented in section 3.2, restricts its sample relative to the headline metric and it additionally serves as the primary physical benchmark in the InSAR volume analysis of section 5.8.

In practice, equation (2) is estimated by applying the two-way within transformation to all variables and running OLS on the demeaned data. Specifically, define $\tilde{y}_{i,t} = y_{i,t} - \bar{y}_{i,\cdot} - \bar{y}_{\cdot,t} + \bar{y}_{\cdot,\cdot}$, where $\bar{y}_{i,\cdot}$ denotes the within-mine mean, $\bar{y}_{\cdot,t}$ the within-quarter mean, and $\bar{y}_{\cdot,\cdot}$ the grand mean. Letting $\tilde{x}_{i,t}$ denote the analogously transformed *NetSignal*, the TWFE estimator is

$$\hat{\beta}^{\text{TWFE}} = \frac{\sum_{i,t} \tilde{x}_{i,t}\, \tilde{y}_{i,t}}{\sum_{i,t} \tilde{x}_{i,t}^2} \tag{3}$$

Equation (3) makes explicit that $\hat{\beta}^{\mathrm{TWFE}}$ is the OLS slope of demeaned production on demeaned net satellite signal, exploiting only the within-mine, within-quarter variation that survives partialling out the fixed effects.

Although TWFE is the workhorse estimator for panel data and the standard headline specification in the alternative-data finance literature (see e.g., Dessaint et al. (2024)), the efficiency of conventional inference depends on the residual structure. We estimate within-mine residual autocorrelation to assess the persistence of idiosyncratic shocks. The mean (median) AR(1) coefficient across mines is approximately 0.46 (0.55), indicating moderate persistence, with substantial heterogeneity across mines. The estimated autocorrelation suggests non-negligible persistence in the idiosyncratic error term. This motivates reporting both TWFE and first-difference specifications, as their relative efficiency depends on the persistence of both the error process and the regressors. We therefore complement equation (2) with the FD specification:

$$\Delta \log \left(\mathrm{Production}_{i,t}\right) = \gamma_t + \beta^{\mathrm{FD}} \cdot \Delta \mathrm{NetSignal}_{i,t} + u_{i,t} \tag{4}$$

where Δ denotes the first difference between consecutive quarters and $u_{i,t} \equiv \varepsilon_{i,t} - \varepsilon_{i,t-1}$. The FD specification differences out the mine fixed effects α_i by construction, retains quarter fixed effects γ_t to absorb common shocks, and identifies β^{FD} from quarter-on-quarter changes in production and signal. Specifically, FD is consistent under the weaker assumption of contemporaneous exogeneity in differences, $\mathbb{E}[\Delta \mathrm{NetSignal}_{i,t}\, u_{i,t}] = 0$, which does not require strict exogeneity over the full sample period. We compute cluster-robust standard errors at the mine level, allowing for arbitrary heteroskedasticity and serial correlation within mines.

4.3 Heterogeneity and dynamic response

The pooled specifications in equations (2) and (4) restrict the response of production to satellite signals to be common across mines and constant within the sample period, and assume that the entire informational content of the signal is realised contemporaneously. Both sets of restrictions are strong. Specifically, the production technologies, ore grades, and operational scales differ substantially across commodities, and the global commodity-price regime that prevailed in 2018–2021 differs materially from the post-2022 environment. Furthermore, a contemporaneous specification rules out delayed responses to operational changes by construction. In this subsection, we relax each restriction in turn.

To assess heterogeneity, we re-estimate the first-difference specification in equation (4) on subsamples defined along four dimensions: commodity class, mine size, sub-period, and satellite data quality. Formally, for each subsample $\mathcal{S}$ we estimate

$$\Delta \log \left(\mathrm{Production}_{i,t}\right) = \gamma_t + \beta_{\mathcal{S}} \cdot \Delta \mathrm{NetSignal}_{i,t} + u_{i,t}, \quad (i,t) \in \mathcal{S} \tag{5}$$

We report heterogeneity in the first-difference specification because differencing addresses by construction the level heterogeneity across commodity classes and mine sizes that complicates the interpretation of TWFE estimates in stratified samples, where the within-subsample fixed effects absorb a more restricted set of variation. The commodity subsamples are constructed by partitioning mines according to their primary commodity (copper, gold, iron ore), as classified in section 3.2. The mine-size subsamples are constructed by a within-commodity median split of average reported quarterly *Commodity Production*, which holds commodity-specific scale conventions constant and isolates relative operational size. The sub-period split partitions the sample at Q1 2022, reflecting the commodity-price regime shift between the pre-pandemic period and the post-2022 environment (Baffes & Nagle, 2022). In addition, the data-quality split classifies mines as high-retention ($\geq 80\%$ of mine-quarters meeting the 30-day cloud-free threshold described in section 3.3) or medium-retention (40%–80%); whereas the other three dimensions test economic heterogeneity, this fourth split tests whether the predictive content of the signal is concentrated where it is most reliably measured. Inference within each subsample is done with cluster-robust standard errors at the mine level, following the same methodology as in section 4.2.

To examine whether the response of production to satellite signals is concentrated contemporaneously or distributed over adjacent quarters, we expand the TWFE specification in equation (2) with leads and lags of the net signal:

$$\log(\text{Production}_{i,t}) = \alpha_i + \gamma_t + \sum_{k=-K}^{+K} \beta_k \cdot \text{NetSignal}_{i,t-k} + \varepsilon_{i,t} \tag{6}$$

where β_0 recovers the contemporaneous effect, β_k for $k > 0$ captures the response of production at t to past signals, and β_k for $k < 0$ tests for anticipation effects. We set $K = 2$. The lead coefficients β_{-1} and β_{-2} provide a falsification test of the strict exogeneity assumption (1): coefficients statistically distinguishable from zero would indicate that future signals carry information about current production conditional on the fixed effects, inconsistent with the identifying assumption that anchors the main specification. Equation (6) delivers both a characterisation of the dynamic response and an empirical check on identification.

4.4 Firm-level earnings predictability

Having established the mine-level satellite-production relationship, we ask whether it transmits into financially relevant information at the firm level. We implement two analyses in sequence. The first tests whether the aggregated NO_2 signal predicts the direction of earnings surprises, establishing the transmission channel from physical extraction to reported financial outcomes. The second asks the more demanding question of whether the satellite-derived forecast contains information incremental to the analyst consensus, testing whether the signal adds predictive content beyond what is already reflected in market expectations.

4.4.1 Predicting earnings surprise direction

We ask whether the aggregated NO_2 signal predicts the direction of earnings surprises at the earnings announcement, and describe in turn the sample restriction, signal aggregation procedure, and binary response model used to test this relationship. A central challenge in aggregating mine-level signals to the corporation level is that most firms in our sample operate more mines than are represented in the satellite panel, introducing a wedge between the observed signal and the firm's true operational activity. To address this, we restrict the analysis to corporations for which at least 50 percent of their mines are represented in the satellite panel, ensuring that the aggregated signal captures a sufficiently large share of firm-level extraction activity to support a directional prediction. This restriction yields a final estimation sample of 126 firm-period observations. For each firm-reporting period, the satellite signal is constructed as the simple average of the NO_2 net BBox signal across all of the firm's sample mines in the corresponding reporting window. Specifically, for quarterly reporters the signal is a single-quarter average, whereas for half-yearly reporters it is the average of the two constituent quarters' signals.

The dependent variable is a binary indicator $\text{Beat}_{f,t}$ equal to one if adjusted reported EPS exceeds the adjusted Bloomberg consensus estimate in period t for firm f, and zero otherwise. The binary nature of the outcome variable motivates the use of a logit specification. Specifically, we estimate:

$$\Pr\left(\text{Beat}_{f,t} = 1\right) = \Lambda\left(\alpha + \beta \cdot \text{NetSignal}_{f,t} + \gamma \cdot \text{HalfYear}_{f,t}\right) \tag{7}$$

where $\Lambda(\cdot)$ denotes the logistic function and $\text{HalfYear}_{f,t}$ is a dummy variable equal to one for firms reporting on a half-yearly basis. The parameter β is interpreted through the corresponding odds ratio $\exp(\hat{\beta})$, which measures the change in odds of beating the consensus associated with a unit change in the net satellite signal, holding all other variables constant. We estimate two variants of equation (7) that differ in the construction of $NetSignal_{f,t}$. The first uses the continuous standardised firm-level NO_2 signal, testing whether the signal enters linearly in the log-odds of beating consensus. For the second variant we define a binary above-median indicator equal to one if the firm's aggregated signal in the reporting period exceeds its own historical median, and zero otherwise. This analysis is motivated by the hypothesis that the satellite-earnings relationship is better characterised as a threshold effect than as a linear function of signal magnitude. The half-yearly reporting dummy is included in both specifications to control for potential differences in beat rates across reporting frequencies; it is not statistically significant across specifications, suggesting limited evidence that reporting frequency drives the results.

4.4.2 Incremental predictive content over analyst consensus

Although equation (7) tests whether the satellite signal predicts the direction of earnings surprises, it does not establish whether the signal contains information that is incremental to the analyst consensus. Specifically, the logit specification mentioned in section 4.4.1 cannot distinguish between two empirically distinct possibilities: that the satellite signal carries information about realized earnings that analysts have not yet incorporated, and that the signal predicts beats merely because it correlates with information already embedded in the consensus forecast. To address this directly, we estimate an encompassing regression of realized adjusted EPS on both the firm-period production forecast generated by the first differences model of Section 4.2 and the Bloomberg consensus estimate. The fitted production forecast is constructed in two steps. First, for each mine-quarter, we compute $\Delta \widehat{\log(\text{Prod}_{i,t})} = \hat{\beta}^{FD} \cdot \Delta\text{NetSignal}_{i,t}$, where $\hat{\beta}^{FD}$ is the slope coefficient from the first-differences specification in equation (4). Second, we aggregate these fitted values to the firm-period level using the same procedure described in Section 4.4.1: for quarterly reporters we take the firm-quarter mean across the firm's sample mines, and for half-yearly reporters we additionally average across the two constituent quarters. Formally, we estimate

$$\text{EPS}_{f,t} = \alpha_f + \beta_1 \cdot \Delta\,\widehat{\log(\text{Prod})}_{f,t} + \beta_2 \cdot \text{Consensus}_{f,t} + \varepsilon_{f,t} \tag{8}$$

where α_f denotes firm fixed effects and standard errors are clustered at the firm level. The fitted forecast enters the specification in its natural units rather than as a standardised variable, which preserves the economic interpretability of the coefficient: β_1 measures the dollar impact on realized EPS per unit change in the satellite-predicted log production. A half-year reporter dummy is not included, in contrast to equation (7), because every firm in the sample reports consistently in either quarterly or half-yearly frequency over the observation window; the half-year indicator is therefore time-invariant within firm and is absorbed by α_f by construction. The parameter of interest is β_1: if the satellite-based production forecast contains predictive content beyond what is already incorporated in the analyst consensus, β_1 should be statistically distinguishable from zero after controlling for the consensus estimate. The coefficient on the consensus, β_2, is expected to be close to unity if the consensus is an unbiased forecast of realized EPS. Equation (8) is estimated on the same high-coverage firm sample as the logit in equation (7), with the full 17-firm panel reported as a robustness check.

4.5 Trading strategy construction

Having established that the satellite signal predicts the direction of earnings surprises for high-coverage firms, we assess whether this predictive relationship is exploitable in a systematic long-short equity strategy around earnings announcements. We implement two strategies in sequence. The first constructs a binary long-short strategy based on the above-median signal and evaluates its announcement-day performance under both in-sample and out-of-sample variants. The second extends this framework by constructing a probability-based strategy using recursively estimated logit fitted values, evaluating both an intuitive and a contrarian implementation across multiple announcement windows and confidence thresholds.

4.5.1 Binary trading strategy

The trading strategy exploits the binary above-median signal constructed in section 4.4. Specifically, for each earnings announcement we take a long position in firms whose pre-announcement NO_2 signal exceeds their own historical median and a short position in firms whose signal falls below it. The position is opened at the announcement date and closed at the one-day return, so that the strategy captures the market's immediate price response to the disclosed earnings rather than any subsequent drift. The analysis is restricted to the six firms with at least 50 percent mine coverage, consistent with the sample restriction in section 4.4, covering 110 announcements in total.

We evaluate two variants of the strategy that differ in how the historical median threshold is formed. Variant A uses each firm's full-sample median as the threshold, which implies that the classification of each announcement as a high- or low-signal event is determined using information from the full observation window. Variant A is an in-sample specification and is reported for descriptive purposes only. Variant B addresses this limitation by replacing the full-sample median with an expanding within-firm median computed from all prior announcements for that firm at the time of each trade. Specifically, at each announcement date t for firm f, the threshold is set equal to the median of all previously observed firm-level signals up to but not including t, ensuring that no future information enters the classification. Variant B therefore constitutes a genuine out-of-sample backtest in the spirit of Katona et al. (2025). A minimum of ten prior observations is required before variant B begins trading, which determines the burn-in period and the first tradeable date.

In line with the small number of firms in the sample, conventional asymptotic inference is unreliable for evaluating strategy performance. We therefore report bootstrap p-values based on firm-block resampling with 5,000 replications, in which entire firm-level return sequences are resampled with replacement. This procedure preserves the within-firm serial dependence structure of announcement returns while allowing non-parametric inference on the pooled mean strategy return. Bootstrap p-values are two-sided, computed as the fraction of bootstrap samples in which the absolute resampled mean exceeds the absolute observed mean, corresponding to a null of zero mean per-trade return.

4.5.2 Out-of-sample logit trading strategy

To evaluate whether the earnings-direction signal documented in section 4.4.1 is exploitable in real time, we extend the baseline trading framework by constructing an out-of-sample probability-based strategy using the fitted values from the logit specification in equation 7. Specifically, for each firm-period observation, the recursively estimated logit model produces a fitted probability

$$\hat{p}_{f,t} \equiv \Pr(\text{Beat}_{f,t} = 1),$$

which is interpreted as the model-implied probability that firm f will report adjusted EPS above the Bloomberg consensus estimate at announcement date t.

To avoid look-ahead bias, probabilities are generated recursively using an expanding-window estimation procedure. Beginning after an initial burn-in period of 20 observations, equation 7 is sequentially re-estimated using only information available prior to each announcement, and a one-step-ahead fitted probability is generated for the subsequent observation. A natural directional implementation would take long positions in firms with high predicted beat probabilities and short positions in firms with low predicted beat probabilities. Alongside this intuitive implementation, we additionally evaluate a contrarian variant in which positions are inverted relative to the predicted earnings surprise direction, capturing the possibility that sophisticated investors partially incorporate and overweight the signal ahead of announcements.

Trades are executed only when the fitted probability deviates sufficiently from the neutral probability of 0.5. Formally, positions are entered when

$$\hat{p}_{f,t} < 0.5 - \tau \quad \text{or} \quad \hat{p}_{f,t} > 0.5 + \tau,$$

where $\tau \in \{0.00, 0.05, 0.10, 0.15\}$ defines the minimum confidence threshold required to initiate a trade. Under the inverted specification, observations satisfying $\hat{p}_{f,t} < 0.5 - \tau$ generate long positions, whereas observations satisfying $\hat{p}_{f,t} > 0.5 + \tau$ generate short positions.

We additionally evaluate multiple announcement windows to assess the temporal concentration of abnormal returns around earnings releases. Specifically, cumulative abnormal returns are computed over the $[-1, +1]$, $[-5, +1]$, and $[-10, +1]$ windows relative to the earnings announcement date. Strategy performance is again evaluated using mean returns, hit rates, Sharpe ratios, and firm-block bootstrap inference with 5,000 replications.

5 Empirical results

This section presents the empirical results of the satellite-production analysis, building on the econometric framework established in section 4. We begin in section 5.1 with descriptive evidence that establishes whether the satellite-production hypothesis has visual support in the data before any formal estimation. Section 5.2 then reports the baseline contemporaneous TWFE estimates, which are positive and directionally consistent but fall short of conventional significance thresholds. Section 5.3 addresses the temporal structure of the relationship, confirming that the signal responds contemporaneously rather than leading or lagging production. Section 5.4 presents the preferred first-differences specification, which yields the headline significant results. Section 5.5 unpacks where the relationship is strongest and where it breaks down, examining commodity type, mine size, sub-period, and data quality. Section 5.6 verifies the stability of the main findings across alternative geometries and signal definitions. Finally, sections 5.7 and 5.8 extend the framework to firm-level earnings surprises, the incremental predictive content of the satellite-derived forecast relative to the analyst consensus, a trading strategy, and a physically independent volume-based cross-check.

5.1 Preliminary evidence

Before turning to the formal regression analysis, we examine the raw within-mine relationship between satellite signals and production to establish whether the hypothesis has descriptive support in the data. Figure 1 presents scatter plots of within-mine demeaned log production against within-mine demeaned net signal for all six signal-outcome combinations, with both the dependent variable and the regressor expressed as deviations from their mine-specific means. This is precisely the variation that the two-way fixed effects and first-differences estimators exploit, and the plots therefore serve as a direct visual pre-check on the identifying variation.

The NO_2 panels reveal a consistently positive within-mine relationship across all three production metrics. The pooled within-mine correlation is $r = +0.041$ for commodity production ($N = 739$), $r = +0.060$ for material moved ($N = 387$), and $r = +0.049$ for ore mined ($N = 352$), with corresponding OLS slopes of 5,230, 4,960, and 5,250 respectively. While these correlations are modest in magnitude, their consistency across all three production metrics is notable given the substantial noise inherent in quarterly satellite retrievals and the heterogeneous operational contexts of the 38 mines contributing to the NO_2 sample. The aerosol index panels present a weaker and less consistent picture. The AI signal loads positively on commodity production ($r = +0.055$, $N = 1{,}050$) and material moved ($r = +0.048$, $N = 597$), but the ore mined panel yields a slightly negative relationship ($r = -0.019$, $N = 522$, slope $= -0.084$). The AI-ore mined combination is the only panel in which the sign runs counter to the hypothesised direction, and the magnitude is negligible. Across all panels, the scatter is wide, with substantial dispersion around the fitted line, reflecting the cross-mine heterogeneity we examine formally in section 5.5.

Taken together, the demeaned scatter plots provide weak but directionally consistent preliminary support for the satellite-production hypothesis for NO_2, and a more mixed picture for the aerosol index. The modest correlations also set realistic expectations for the regression results that follow: the formal estimators will sharpen identification by adding time fixed effects and, in the preferred specification, first-differencing, but they operate on the same underlying within-mine variation visible here.

Figure 2 documents the cross-mine heterogeneity underlying the pooled correlations. For the NO_2 signal, within-mine time-series correlations between net signal and log commodity production range from $r = +0.733$ at Chapada to $r = -0.326$ at Quellaveco across the 32 mines with at least eight joint observations, with a median of $r = +0.261$ and a mean of $r = +0.193$. Nineteen of the 32 mines exhibit $r > +0.20$ and five exceed $r > +0.40$, while only four mines show $r < -0.20$: Iron Bridge, El Soldado, Sabodala-Massawa, and Quellaveco. The AI signal exhibits a considerably more dispersed and symmetric distribution across the 38 mines with sufficient joint observations, with a median within-mine correlation of only $r = +0.068$ and a mean of $r = +0.063$; only 10 of the 38 mines exceed $r = +0.20$, and the mines driving the strongest positive AI correlations, Caserones ($r = +0.538$) and Gruyere ($r = +0.488$), are largely distinct from those driving the NO_2 signal, consistent with the low cross-signal correlation of $r = 0.217$ documented in section 3.3. This heterogeneity motivates the commodity and data-quality splits examined formally in section 5.5.

5.2 Main results: contemporaneous TWFE

Table 1 reports the baseline two-way fixed effects estimates for all signal-outcome combinations under the primary bbox-control geometry. For commodity production, the NO_2 net signal yields $\hat{\beta} = +7{,}036$ ($SE = 4{,}631$, $p = 0.129$, $N = 739$) and the aerosol index yields $\hat{\beta} = +0.303$ ($SE = 0.207$, $p = 0.143$, $N = 1{,}050$). Both coefficients are positive and directionally consistent with the hypothesised diesel combustion and dust mechanisms, but neither reaches conventional significance thresholds. The within-R^2 is negligible across all specifications, indicating that the net signal explains little of the residual within-mine quarterly variation once mine and time fixed effects are absorbed, and that quarterly variation in production is dominated by other factors with satellite signals capturing a small but directionally consistent component.

The positive sign for NO_2 holds across all three production metrics. For material moved the estimate is $\hat{\beta} = +3{,}409$ ($p = 0.516$, $N = 387$) and for ore mined $\hat{\beta} = +7{,}538$ ($p = 0.212$, $N = 352$). The aerosol index is similarly positive for commodity production and material moved, but loads negatively on ore mined ($\hat{\beta} = -0.079$, $p = 0.616$), the only sign reversal across the full baseline table, and one of negligible economic magnitude. When both signals are entered jointly in Specification 2, the NO_2 and AI coefficients remain broadly stable across all three dependent variables; $\hat{\beta}_{NO_2} = +7{,}188$ and $\hat{\beta}_{AI} = +0.237$ for commodity production, confirming that the two signals carry largely independent information, consistent with their low mutual correlation of $r = 0.217$. Taken together, the contemporaneous TWFE estimates establish a positive and directionally consistent baseline relationship between satellite-derived signals and mine-level production, with significance emerging in specific sub-samples and under the preferred first-differences specification documented in sections 5.3–5.6.

5.3 Lead-lag structure

To assess whether the satellite signal leads or lags reported production, and to test for reverse causality, we augment the baseline specification with two leads and two lags of the net signal, estimating the response of log commodity production to the net signal at horizons $t-2$ through $t+2$. Table 14 reports the results for both the NO_2 and aerosol index signals.

The pattern across all five horizons is notably flat for both signals. For NO_2, no coefficient at any horizon approaches conventional significance, with p-values ranging from 0.206 at the contemporaneous horizon to 0.856 at $t-1$ ($N = 648$). The contemporaneous coefficient ($\hat{\beta}_t = +4{,}099$, $SE = 3{,}240$, $p = 0.206$) is the largest in absolute magnitude among the five horizons, but the differences across horizons are economically modest and statistically indistinguishable. The aerosol index exhibits a similarly flat profile, with all five coefficients insignificant and no systematic pattern across the lag structure (p-values ranging from 0.196 to 0.924, $N = 920$).

Two methodologically important conclusions follow from this pattern. First, the absence of significant lead coefficients at $t + 1$ and $t + 2$ provides no evidence of reverse causality: reported production does not appear to predict future satellite signals, which would be the case if mine operators anticipated and disclosed production in advance of the operational activity detected by the satellite. Second, the absence of significant pre-trend coefficients at $t - 1$ and $t - 2$ is consistent with the satellite signal responding to contemporaneous operational activity rather than to anticipatory changes in mine behaviour. Together these findings validate the contemporaneous specification adopted throughout the analysis and support interpreting the satellite-production relationship as operationally driven rather than artefactual.

5.4 Preferred specification: first-differences

As documented in section 4.2, the levels residuals from the TWFE specification exhibit moderate positive serial correlation, motivating first-differencing as an additional model. By construction, first differences absorb all mine-specific linear trends and restrict identification to quarter-on-quarter co-movement between changes in the satellite signal and changes in production, providing a more credible test of the contemporaneous satellite-production relationship than the levels specification.

The first-differences analysis focuses on commodity production and ore mined, the two metrics for which the identifying assumption underlying the FD estimator is most credible. Material moved, whilst reported in the TWFE baseline for completeness, is excluded from the preferred specification on two grounds. First, its disclosure coverage is substantially lower than the other metrics, and first-differencing further reduces the effective sample by eliminating the first observation for each mine, leaving an unbalanced and potentially unrepresentative panel. Second, and more fundamentally, quarterly changes in total material moved reflect not only active ore extraction but also irregular waste-stripping campaigns and overburden-removal schedules that operate on multi-quarter planning cycles entirely disconnected from the within-quarter haulage activity captured by the satellite signal. This introduces a systematic wedge between quarter-on-quarter variation in material moved and the contemporaneous diesel emissions detected by Sentinel-5P, violating the implicit assumption that the differenced dependent variable and the differenced regressor co-move within the same quarter. Ore mined and commodity production are both more tightly linked to continuous operational intensity and are therefore better suited to the first-differences estimator.

Table 2 reports the first-differences estimates for both dependent variables and both signals. The headline finding is the commodity production specification: $\Delta \log(\text{Commodity Production})$ regressed on ΔNO_2 yields $\hat{\beta} = +2{,}690$ ($SE = 1{,}030$, $p = 0.009$, $N = 701$), significant at the one percent level. This is the strongest and most credible result in the full analysis. The diesel combustion mechanism and blasting cycles predicts precisely this relationship: quarters in which mines collectively extract and process more output require more haulage cycles and generate proportionally more tropospheric NO_2, and this co-movement is detectable even after removing all mine-specific trends through first-differencing. The ore mined specification yields a positive and directionally consistent estimate of $\hat{\beta} = +5{,}632$ ($SE = 4{,}550$, $p = 0.216$, $N = 330$), but does not reach conventional significance thresholds, reflecting the substantially smaller sample available for this metric given the limited disclosure coverage documented in section 3.2.

The aerosol index yields no significant result under first differences for either dependent variable. For commodity production $\hat{\beta} = +0.134$ ($p = 0.210$) and for ore mined $\hat{\beta} = +0.006$ ($p = 0.964$), the latter estimate being indistinguishable from zero. This sharp divergence between the NO_2 and aerosol index results under first differences carries a clear interpretation: the diesel combustion channel, captured by tropospheric NO_2, is the operative mechanism linking satellite observations to extraction activity, while the dust and absorbing aerosol channel does not contribute incremental information about production variation at quarterly frequency once persistent mine-level trends are removed. The aerosol index may reflect geological and meteorological variation that co-moves with production in levels but does not respond sensitively to quarter-on-quarter changes in operational intensity. This is consistent with evidence documented in section 2 that boundary-layer stability and meteorological conditions constitute the dominant drivers of observed aerosol concentrations at mine sites, such that the aerosol signal reflects atmospheric state as much as operational intensity.

5.5 Heterogeneity analysis

We examine heterogeneity in the first-differences satellite-production relationship across four dimensions: commodity type, mine size, sub-period, and satellite data quality, with results reported in table 3. Throughout this section the dependent variable is log commodity production or log ore mined, the signal is the NO_2 net bbox mean unless stated otherwise, and all specifications include quarter fixed effects with mine-clustered standard errors. The aerosol index results are reported for completeness but are uniformly insignificant across all heterogeneity dimensions, consistent with the pooled finding in section 5.4.

The pooled NO_2 result masks substantial commodity-level heterogeneity (Panel A). The strongest commodity-level finding is for iron ore mines, where the first-differences estimate for commodity production is $\hat{\beta} = +10{,}774$ ($SE = 4{,}172$, $p = 0.010$, $N = 162$), significant at the one percent level. This result is consistent with the physical mechanism: iron ore operations in our sample are dominated by large-scale Pilbara and South African open-pit mines involving enormous diesel-powered truck fleets operating across arid, low-cloud environments where NO_2 signals are most reliably measured. First-differencing removes the level differences that obscure this relationship in the TWFE specification, revealing a robust contemporaneous link between extraction intensity and atmospheric emissions. Copper mines yield a marginally significant estimate for commodity production ($\hat{\beta} = +2{,}627$, $SE = 1{,}387$, $p = 0.058$, $N = 248$), consistent with the continuous high-volume haulage that characterises large open-pit porphyry deposits. Gold mines, by contrast, yield a negative and insignificant coefficient for commodity production ($\hat{\beta} = -508$, $p = 0.893$) under first differences. This negative sign, consistent across both the TWFE and FD specifications for gold, reflects the weaker operational link between gold output, which is sensitive to ore grade and processing efficiency rather than raw material throughput, and the haulage-driven NO_2 signal captured by Sentinel-5P.

Turning to mine size, we classify mines as large or small based on a within-commodity median split of average quarterly production (Panel B). For commodity production, neither size class produces a significant first-differences estimate, with large mines yielding $\hat{\beta} = +2{,}115$ ($p = 0.139$, $N = 363$) and small mines $\hat{\beta} = +1{,}211$ ($p = 0.702$, $N = 338$). The ore mined specification tells a more differentiated story: small mines yield a significant positive estimate of $\hat{\beta} = +30{,}657$ ($SE = 14{,}974$, $p = 0.041$, $N = 191$), while large mines produce a positive but insignificant coefficient ($\hat{\beta} = +1{,}443$, $p = 0.239$, $N = 139$). The substantially larger coefficient for small mines in the ore mined specification reflects two compounding mechanisms. First, small mines typically operate simpler, more uniform extraction cycles in which diesel consumption is more directly and proportionally linked to ore throughput, whereas large mines run complex multi-bench, multi-equipment operations with auxiliary activities including stockpile management, in-pit crushing, and secondary haulage that generate NO_2 without directly tracking quarterly ore extraction. Second, the first-differences estimator identifies from quarter-on-quarter variation, and small mines exhibit greater relative fluctuations in ore mined as a share of their baseline capacity, making the NO_2 response to changes in operational intensity more detectable against the background signal. For large mines, whose production is deliberately smoothed across quarters through pre-stripping and blending strategies, the within-mine variation exploited by the FD estimator is smaller, attenuating the estimated coefficient even where the underlying physical relationship is equally strong.

Splitting the sample at 2022, the pre-2022 sub-period yields a significant first-differences estimate for commodity production of $\hat{\beta} = +5{,}330$ ($SE = 2{,}123$, $p = 0.012$, $N = 284$), while the post-2022 estimate is smaller and insignificant ($\hat{\beta} = +1{,}207$, $p = 0.564$, $N = 387$) (Panel C). This pattern suggests that the satellite signal was more tightly coupled to production in the earlier part of the observation window, potentially reflecting the more stable operational conditions of the pre-pandemic period relative to the post-2022 years which saw significant disruptions to mine scheduling and supply chains across the sector. For ore mined, both sub-period estimates are positive and of similar magnitude, $\hat{\beta} = +4{,}915$ ($p = 0.254$, $N = 111$) pre-2022 and $\hat{\beta} = +5{,}235$ ($p = 0.576$, $N = 203$) post-2022, but neither reaches significance, likely reflecting the substantially smaller samples available for this metric.

The final heterogeneity dimension is satellite data quality, measured by the share of quarters in which a mine has at least 30 cloud-free NO_2 observations (Panel D). We classify mines as high-retention ($\geq 80\%$ of quarters meeting the threshold) and medium-retention (40–80%). This split is not an economic heterogeneity dimension but a measurement-quality check: if the satellite-production relationship is genuine, it should be strongest precisely where the signal is most reliably measured. The results are strongly consistent with this prediction. For commodity production, high-retention mines yield $\hat{\beta} = +3{,}644$ ($SE = 1{,}618$, $p = 0.024$, $N = 538$), significant at the five percent level, while medium-retention mines yield $\hat{\beta} = +128$ ($p = 0.967$, $N = 131$), indistinguishable from zero. The contrast between the two groups is now sharper than under the previous specification, with the medium-retention coefficient collapsing to near zero. This finding has a direct practical implication: satellite-based production monitoring with NO_2 is most viable at mines in low-cloud environments, predominantly arid and semi-arid climates such as the Atacama Desert, the Sahel, and the Australian Pilbara, and least reliable at mines in persistently cloudy tropical or sub-arctic environments.

5.6 Robustness checks

Table 4 and figure 3 report a battery of ten alternative specifications for the headline NO_2 regressions across both dependent variables. The battery varies four dimensions of the baseline specification: the satellite geometry (bbox-control versus polygon-control on a matched sample of mines with sufficient polygon coverage), the daily cloud threshold (30 percent versus 50 percent), the standard error clustering level (mine, firm, and two-way mine-by-quarter), and the fixed-effects structure (year-quarter fixed effects versus commodity-by-quarter fixed effects). Together these ten specifications test whether the headline findings are sensitive to the specific methodological choices documented in section 3 and section 4.

For the headline result of interest, $\Delta \log(\text{Commodity Production})$ regressed on ΔNO_2, the coefficient is positive across all ten specifications without exception. Statistical significance is concentrated in the first-differences variants: all four FD robustness specifications are significant at the five percent level or below. Specifically, the FD polygon specification yields $\hat{\beta} = +3{,}659$ ($p < 0.001$), the FD firm-clustered specification yields $\hat{\beta} = +2{,}690$ ($p < 0.001$), and the FD commodity-by-quarter fixed effects specification yields $\hat{\beta} = +2{,}671$ ($p = 0.034$). This pattern is precisely what one would expect given the moderate residual serial correlation documented in section 4: the TWFE-levels estimator is directionally consistent but statistically imprecise, while the first-differences estimator removes mine-specific trends and is both consistent and precise. The TWFE specifications are uniformly positive across all six variants but reach conventional significance only rarely, with all specifications in the range $p = 0.108$ to $p = 0.773$. We interpret the robustness battery as strongly supporting the directional finding under TWFE and the statistical finding under FD.

For the ore mined outcome, the directional consistency is equally strong: all ten specifications return positive NO_2 coefficients, confirming the result is not a sign artefact of the headline specification. Statistical significance is less pervasive than for commodity production, which is expected given the substantially smaller sample for this metric. Two TWFE variants reach significance: firm-clustered ($\hat{\beta} = +7{,}538$, $p = 0.034$) and commodity-by-quarter fixed effects ($\hat{\beta} = +9{,}176$, $p = 0.038$), suggesting the ore mined relationship is detectable when the standard error structure accounts for firm-level rather than mine-level clustering. The FD variants are directionally consistent but fall short of conventional thresholds, consistent with the reduced statistical power arising from the smaller ore mined sample.

The aerosol index presents a clean and consistent null result. Across all ten specifications and both dependent variables, not a single AI coefficient reaches the ten percent significance threshold. For commodity production, all ten coefficients are positive and clustered tightly in the range +0.10 to +0.30, statistically indistinguishable from zero throughout. For ore mined, the AI coefficient is slightly negative in the headline specification and mixed in sign across the robustness variants, with p-values ranging from 0.541 to 0.965. This is a genuine null result rather than a statistical power problem: the AI signal is consistently measured with sufficient precision to reject economically meaningful positive effects, and the absence of significance is stable across all ten specification variants. Taken together, the robustness battery confirms the two headline conclusions of the empirical analysis: the NO_2 first-differences relationship with commodity production is robust to all methodological variants tested, and the aerosol index contributes no incremental predictive content under any specification.

5.7 Financial applications

Having established that satellite-derived NO_2 signals contain information about mine-level production, this section examines whether this physical relationship translates into financially exploitable information. We first ask whether satellite signals that predict mine-level production also predict the direction of firm-level earnings surprises, establishing the transmission channel from physical extraction activity to reported financial fundamentals. We then ask whether the satellite-derived production forecast contains information about realised EPS that is incremental to the analyst consensus, testing whether the signal adds predictive content beyond what is already reflected in market expectations. Finally, we evaluate whether this predictive relationship is exploitable in a systematic trading strategy, providing a direct assessment of the alternative data value of satellite-based mining monitoring.

5.7.1 Earnings surprise prediction

We examine whether the satellite-derived NO_2 signal that predicts mine-level production also predicts whether a firm beats its consensus EPS estimate at the earnings announcement. The logit specifications are described in section 4.4; the final estimation sample covers 126 firm-period observations. Results are reported in table 5. The first specification uses the continuous standardised NO_2 signal as the predictor. The estimated coefficient is $\hat{\beta} = +0.392$ ($SE = 0.213$, $z = 1.842$, $p = 0.065$), marginally significant at the ten percent level, with a corresponding odds ratio of 1.48. A one standard deviation increase in the firm's pre-announcement NO_2 signal is thus associated with 48 percent higher odds of beating the consensus estimate. The overall model fit is modest (Pseudo $R^2 = 0.024$, LLR $p = 0.125$), consistent with the signal capturing a real but noisy component of the information set relevant to earnings outcomes.

The second specification replaces the continuous signal with a binary indicator equal to one if the firm's average NO_2 signal in the reporting period exceeds its own historical median, and zero otherwise. The estimated coefficient is $\hat{\beta} = +0.936$ ($SE = 0.394$, $z = 2.374$, $p = 0.018$), significant at the five percent level, with an odds ratio of 2.55. Firms whose satellite signal exceeds their own historical median in the quarter preceding an earnings announcement are thus more than twice as likely to beat the consensus estimate. The overall model fit improves relative to the continuous specification (Pseudo $R^2 = 0.036$, LLR $p = 0.042$), with the full model now significant at conventional thresholds. The stronger result under the binary above-median formulation suggests that the satellite-earnings relationship is better characterised as a threshold effect than as a linear function of signal magnitude. Both specifications point in the same direction: pre-announcement NO_2 signals contain statistically detectable information about the direction of earnings surprises, consistent with the physical production relationship documented in section 5.4 transmitting through to reported financial outcomes.

As a firm-level illustration of the mechanism, we examine Teck Resources in detail. Teck represents the most favourable configuration in our sample for satellite-based earnings prediction for three compounding reasons. First, three of its mines are represented in our panel, accounting for approximately 75 percent of its operational portfolio. Second, two of the three mines, Quebrada Blanca and Carmen de Andacollo, are located in the arid Chilean Atacama and achieve near-perfect NO_2 data retention of 97 and 100 percent respectively, placing them in the high-retention category where the relationship is most reliably detectable, as documented in Panel D of table 3. Third, all three mines extract copper, for which the first-differences estimates are consistently positive and marginally significant. Whilst iron ore emerges as the strongest commodity-level result in the pooled heterogeneity analysis, no iron ore operator in our sample achieves comparable mine coverage at the firm level, making Teck the most complete and internally consistent firm-level test of the framework.

Figure 4 plots Teck's quarterly NO_2 net signal alongside its EPS surprise from Q2 2018 through Q4 2025. The within-firm correlation is $r = +0.52$, substantially above the pooled average. Periods in which Teck's aggregated NO_2 signal exceeds its historical median are systematically associated with positive earnings surprises. A notable exception is the 2022 transition period, during which Quebrada Blanca suspended active extraction as part of its QB1 to QB2 phase transition, suppressing the signal in a way that reflects a structural operational change rather than a genuine decline in throughput. Excluding this window, the directional accuracy of the signal is consistent throughout the observation period. Taken together, the Teck case study illustrates both the potential and the boundary conditions of satellite-based earnings prediction at the firm level.

5.7.2 Encompassing regression: incremental predictive content

The logit analysis in section 5.7.1 establishes that the satellite signal predicts the direction of earnings surprises, but it cannot address a more demanding question: does the satellite-derived production forecast contain information about realised EPS that is incremental to what analysts have already incorporated into the consensus estimate? To test this, we estimate the encompassing regression in equation (8), regressing realised adjusted EPS on both the fitted production forecast $\Delta \widehat{\log(\text{Prod})}_{f,t}$ and the Bloomberg consensus estimate, with firm fixed effects and standard errors clustered at the firm level. Results are reported in table 6.

Column (2) confirms that the analyst consensus is a strong and effectively unbiased predictor of realised EPS: the consensus coefficient is $\hat{\beta}_2 = +1.011$ ($SE = 0.056$, $p < 0.001$) and the within-R^2 reaches 0.857, indicating that the consensus alone explains the vast majority of within-firm variation in adjusted EPS. Column (1) shows that the satellite-derived production forecast carries no significant predictive power in isolation ($\hat{\beta}_1 = +2.732$, $p = 0.507$), which is unsurprising given that the consensus already distils a rich information set, and the fitted forecast is a coarse, mine-level aggregate of a single atmospheric signal.

The key result is in column (3). When both predictors enter jointly, the satellite forecast becomes significant at the five percent level ($\hat{\beta}_1 = +11.759$, $SE = 5.827$, $p = 0.047$) while the consensus coefficient is essentially unchanged at $+1.015$ ($p < 0.001$). The stability of $\hat{\beta}_2$ across columns (2) and (3) implies that the satellite signal is not collinear with the consensus but instead contributes orthogonal predictive content. The within-R^2 rises from 0.857 to 0.862, a modest increment consistent with the signal capturing a real but small share of the residual variation left unexplained by the consensus. Together, these results indicate that the satellite-derived production forecast contains information about realised EPS that analysts have not yet incorporated, satisfying the incremental predictability criterion at the firm level.

The robustness panel, estimated on the full 17-firm sample, shows that the satellite forecast is no longer significant in the encompassing specification ($\hat{\beta}_1 = +2.261$, $p = 0.387$), while the consensus remains highly significant and close to unity throughout. This contrast between the high-coverage and full samples mirrors the pattern documented in section 5.7.1 and the heterogeneity results of section 5.5: the informational content of the satellite signal is concentrated precisely where mine coverage is most complete and signal quality is highest. For firms where satellite observations capture only a minority of operational activity, the aggregated forecast is too noisy to add to the consensus, and the encompassing coefficient collapses to insignificance. The finding therefore reinforces the practical guidance for practitioners: the satellite signal offers incremental earnings predictability for firms with high mine coverage in low-cloud environments, and not for the broader universe of mining operators.

5.7.3 Trading strategy

Having established that the NO_2 signal predicts the direction of earnings surprises for high-coverage firms and contains incremental predictive content over the analyst consensus (sections 5.7.1 and 5.7.2), we ask whether this translates into exploitable announcement-day returns. The sample covers the six high-coverage firms across 110 announcements, with Variant B constituting the genuine out-of-sample test.

Both variants deliver returns indistinguishable from zero, as reported in table 7. Variant A generates a mean per-trade return of -0.224 percent ($p = 0.606$) and Variant B generates -0.045 percent ($p = 0.924$, $N = 100$ trades, Sharpe $= -0.010$). The bootstrap 95 percent confidence interval is $[-1.420\%, +1.036\%]$ with a two-sided p-value of 0.967. The absence of tradeable returns is puzzling given the significant logit result in section 5.7.1. Table 8 resolves this contradiction. The signal is informative: high-signal firms beat the consensus in 69.4

percent of cases versus 39.2 percent for low-signal firms. Beats are rewarded at +1.411 percent against −0.311 percent for misses. The strategy fails because when a high-signal firm misses, the market punishes it by −1.334 percent, nearly as large as the +1.433 percent reward for a confirmed beat. The expected long-leg return of +0.586 percent is virtually identical to the short-leg return of +0.650 percent, and the spread collapses to −0.064 percent.

One plausible interpretation is that the satellite signal is symmetrically incorporated into market expectations before the announcement, such that any alpha must be captured pre-announcement rather than on the day itself. The per-firm decomposition shows Kinross Gold (+1.43%, $N = 28$) and Fortescue (+1.07%, $N = 13$) contributing positively while Capstone Copper is the largest drag at −2.65% across 16 announcements. Taken together, the trading strategy analysis confirms that the NO_2 signal is informative about earnings direction but that this information appears to be at least partially priced in before the announcement window opens, motivating the pre-announcement strategy examined in section 5.7.4.

5.7.4 Pre-announcement trading strategy and contrarian returns

The announcement-day results in section 5.7.3 suggest that any alpha must be captured pre-announcement. As a first step, we implement the intuitive strategy of going long on high-NO_2 firms and short on low-NO_2 firms in the days preceding announcements. This strategy yields negative and marginally significant returns across all pre-announcement windows, as reported in table 15, consistent with the satellite signal being not merely priced in but overweighted by sophisticated investors ahead of announcements. This motivates a contrarian strategy that fades the signal direction, which we evaluate using the probability-threshold framework described in section 4.5.2.

Results are reported in table 9. The $[-1, +1]$ window delivers the strongest and most consistent results. At $\tau = 0.00$, the mean per-trade mining-adjusted return is +1.275 percent ($p = 0.051$, $N = 100$, Sharpe = 0.215, hit rate = 60.0 percent). Tightening the threshold to $\tau = 0.05$ improves performance to +1.722 percent per trade ($p = 0.039$, $N = 68$, Sharpe = 0.289, bootstrap 95 percent CI $[+0.05\%, +3.23\%]$), with the improvement in both magnitude and significance reflecting the filtering of low-conviction predictions. At higher thresholds the point estimates remain positive but statistical significance fades as the number of trades declines. The $[-5, +1]$ and $[-10, +1]$ windows produce positive but uniformly insignificant results across all threshold levels, consistent with noise dominating signal over longer pre-announcement horizons.

Figure 5 plots the simple cumulative strategy return over the out-of-sample period at $\tau = 0.00$. The $[-1, +1]$ window accumulates +127.5 percent over the sample, with gains concentrated in the 2021 to 2024 period and the return path broadly positive throughout. The $[-5, +1]$ window reaches +74.3 percent with a similar upward trajectory, while the $[-10, +1]$ window ends at +14.2 percent with a flatter and more volatile path, consistent with the weaker statistical results for that window. Figure 6 confirms that positive mean returns are robust across all threshold levels in the $[-1, +1]$ window, with the $\tau = 0.05$ bar the only one reaching conventional significance.

The pattern is consistent with an overreaction mechanism. As documented in the literature review (section 2.1), gains from satellite-based alternative data accrue disproportionately to sophisticated investors, consistent with active pre-announcement trading on a freely discoverable signal. If this trading overshoots the signal's true predictive content, high-NO_2 firms enter announcements with inflated expectations, generating a contrarian return premium. This interpretation is reinforced by the encompassing regression in section 5.7.2: sell-side analysts underreact to the satellite signal while sophisticated buy-side investors appear to overreact, creating a wedge that the pre-announcement contrarian strategy exploits. Together, the announcement-day and pre-announcement results paint a coherent picture of how a freely available alternative data signal propagates through the price discovery process.

5.8 Extension: InSAR volume estimates

The InSAR volume analysis serves as a physically independent cross-check on the NO_2-based production relationship; its scope and inferential limitations are described in sections 3.5. Within-mine correlations between InSAR-derived volume and log material moved are positive for five of the six mines, ranging from $r = +0.016$ at Antamina to $r = +0.501$ at Damang, with Tarkwa the sole mine exhibiting a negative correlation ($r = -0.258$). Figure 8 reports the per-mine first-difference coefficients, confirming that four of the six mines yield positive estimates, with Pueblo Viejo and Damang showing the strongest positive relationships and wide confidence intervals throughout reflecting the limited within-mine time series available at each site.

Table 16 reports the regression results across specifications that progressively add controls and robustness checks, with MM denoting material moved specifications and CP denoting commodity production specifications. In MM-Joint, the primary material moved specification, the volume coefficient is positive ($\hat{\beta} = +0.069$) but insignificant under wild cluster bootstrap inference ($p_{\text{boot}} = 0.570$). Adding a lagged volume term in MM-Lag does not materially change the conclusion: the contemporaneous coefficient is $\hat{\beta} = +0.079$ ($p_{\text{boot}} = 0.525$) and the lagged coefficient is $\hat{\beta} = +0.068$ ($p_{\text{boot}} = 0.430$). All material moved specifications are positive in sign and consistent in magnitude, but none approach conventional significance thresholds under bootstrap inference.

The commodity production specifications tell a more informative story. In CP-Joint, the InSAR volume coefficient is positive and significant at the five percent level under bootstrap inference ($\hat{\beta} = +0.037$, $p_{\text{boot}} = 0.016$), suggesting that quarterly InSAR-derived extraction volumes contain detectable information about downstream commodity output. Adding a lagged volume term in CP-Lag yields a contemporaneous coefficient of $\hat{\beta} = +0.033$ ($p_{\text{boot}} = 0.061$), marginally significant at the ten percent level, while the lagged term is indistinguishable from zero ($p_{\text{boot}} = 0.840$). The gold-only robustness specification CP-Gold yields a positive volume coefficient of $\hat{\beta} = +0.039$ that reaches conventional significance asymptotically ($p_{\text{asym}} = 0.004$) but collapses to $p_{\text{boot}} = 0.118$ under bootstrap inference with five clusters, underscoring the sensitivity of the result to the cluster count.

The consistently positive directional signs across all material moved specifications, combined with the commodity production result reaching significance at the five percent level under bootstrap inference in CP-Joint, are encouraging: they suggest the physical mechanism is present and detectable even in this constrained subsample. Full validation awaits a larger dataset. Three factors limit what the current analysis can establish. First, InSAR displacement processing is computationally intensive, restricting the sample to six mines and precluding the panel depth needed for reliable cluster inference. Second, the 92.39 m $\times$ 81.66 m pixel grid averages extraction activity across a ground area that often exceeds individual bench widths, attenuating the quarter-on-quarter volume signal relative to what finer-resolution products could recover. Third, surface coherence failures at several candidate sites, described in section 3.5, further constrain the sample. As Sentinel-1 archives deepen and higher-resolution InSAR products become accessible, radar-based volumetric sensing is a promising complement to the NO_2 channel.

6 Summary and concluding remarks

This thesis set out to assess whether satellite-derived atmospheric signals from Sentinel-5P can serve as a real-time proxy for open-pit mining production and whether this physical relationship transmits into financially exploitable information at the firm level. The motivation is straightforward: mining production is a primary driver of cash flows for resource firms, yet it is disclosed on a quarterly basis with material lag between physical extraction and corporate reporting. Conventional alternative data cannot bridge this gap. Satellite-derived atmospheric signals, by contrast, are observable in near real time, openly accessible, and physically grounded in the diesel combustion and dust generation that characterise open-pit extraction.

Using a hand-collected panel of 39 mines from 17 major mining corporations spanning Q2 2018 to Q4 2025, we test whether tropospheric NO_2 and the Aerosol Index, each net of a spatially matched control area, predict reported production under a two-way fixed effects and first-differences framework. The headline finding is that tropospheric NO_2 contains statistically detectable information about within-mine production variation under the preferred first-differences specification ($\hat{\beta} = +2{,}690$, $p = 0.009$, $N = 701$), robust across all ten alternative specifications in the robustness battery and significant in two of these at the one percent level and one at the 5 percent level. The TWFE-levels estimator confirms the positive directional relationship across all three production metrics and all six signal-outcome combinations, though it falls short of conventional significance thresholds, consistent with the moderate residual serial correlation that motivates the first-differences estimator as the preferred specification. The lead-lag analysis confirms that the relationship is strictly contemporaneous, with no evidence of reverse causality or anticipation effects at any horizon from $t - 2$ to $t + 2$, validating the contemporaneous design and supporting an operationally driven interpretation of the signal.

Heterogeneity analysis reveals that the relationship is not uniform across the sample. Iron ore mines emerge as the strongest commodity subgroup ($\hat{\beta} = +10{,}774$, $p = 0.010$), consistent with the large-scale arid-environment Pilbara and South African operations that dominate this subsample. Copper mines yield a marginally significant result ($p = 0.058$), while gold mines produce a negative and insignificant coefficient, reflecting the weaker link between gold output and physical haulage intensity. The data quality split confirms that the relationship is strongest precisely where cloud-free satellite coverage is most reliable: high-retention mines yield $\hat{\beta} = +3{,}644$ ($p = 0.024$) while medium-retention mines collapse to near zero, providing a clear practical criterion for practitioners. Furthermore, investigating the difference in predictive power between size of mines reveals that the relationship is strongest for smaller mines with ore mined. The aerosol index produces a clean null result across all specifications, confirming that the NO_2 channel rather than the dust channel is the operative mechanism at quarterly frequency.

At the firm level, above-median NO_2 signals are associated with more than twice the odds of beating the analyst consensus EPS estimate for high-coverage firms (odds ratio 2.55, $p = 0.018$), establishing a statistically significant transmission channel from physical extraction activity to reported financial outcomes. An encompassing regression further confirms that the satellite-derived production forecast contains predictive content incremental to the Bloomberg analyst consensus: when both the fitted production forecast and the consensus estimate enter jointly, the satellite forecast is significant at the five percent level while the consensus coefficient remains close to unity and unchanged, indicating that the two predictors carry orthogonal information about realised earnings. This incremental content is concentrated in high-coverage firms and collapses to insignificance in the full sample, consistent with the signal quality heterogeneity documented in the production analysis. A systematic long-short trading strategy based on this signal yields announcement-day returns indistinguishable from zero. A 2x2 decomposition reveals that this null is not due to a weak signal but to symmetric market pricing of expectations on both legs: high-signal firms enter announcements with elevated expectations such that beats generate muted rewards while misses are severely punished, leaving the long-short spread near zero. This is consistent with satellite information being at least partially incorporated into equity prices before the announcement window opens. A pre-announcement contrarian strategy that fades the signal direction generates positive mining-adjusted returns of +1.72 percent per trade ($p = 0.039$) over the $[-1, +1]$ window, consistent with sophisticated investors overreacting to a freely discoverable signal ahead of announcements and thereby creating a reversal premium that the contrarian strategy captures. Together, the announcement-day and pre-announcement results paint a coherent picture of how a freely available alternative data signal propagates through the price discovery process: the signal is informative, partially incorporated pre-announcement by sophisticated investors, and overweighted to a degree that generates a contrarian opportunity in the immediate announcement window. The InSAR volume analysis yields null results under bootstrap inference for the material moved specifications, though the commodity production specification reaches significance at the five percent level under bootstrap inference ($\hat{\beta} = +0.037$, $p_{\text{boot}} = 0.016$), suggesting that quarterly InSAR-derived extraction volumes contain detectable information about downstream commodity output even where the direct physical comparison with material moved falls short of significance. Directional signs in the material moved specifications are consistently positive, reinforcing that the physical mechanism is present but not yet precisely estimable given the current dataset size. The restricted sample reflects three compounding constraints: the absence of material moved disclosures for the majority of mines in the main panel, gaps in Sentinel-1 acquisition coverage and instrument-level data quality failures at several sites, and persistent surface coherence issues at mines located in areas of low coherence that cause the SBAS algorithm to fail to converge on stable displacement estimates.

This thesis makes four contributions. First, it provides the first systematic empirical assessment of satellite-derived atmospheric signals as a real-time proxy for mining production, extending the alternative data finance literature from consumer-facing sectors to industrial commodity extraction. Second, it documents a set of methodological choices specific to Sentinel-5P industrial monitoring, including appropriate spatial aggregation relative to mine footprints and the construction of control-area subtraction to isolate mining-specific atmospheric signatures. Third, it identifies the mine and commodity types for which the signal is most informative, providing practical guidance for practitioners seeking to deploy this signal class in investment practice. Fourth, it provides evidence on how a freely available alternative data signal is absorbed into equity prices: the encompassing regression establishes that the signal contains information incremental to sell-side analysts, the announcement-day null suggests at least partial pre-announcement incorporation by sophisticated investors, and the contrarian pre-announcement returns are consistent with that incorporation overshooting the signal's true predictive content.

Several limitations are worth acknowledging. The cross-sectional dimension of the sample is modest at 39 mines and 17 firms, and the firm-level EPS and trading analyses are further constrained to six high-coverage firms, limiting the statistical power of the financial applications. The InSAR volume dataset remains preliminary, covering only a small subset of mines over a shorter observation window than the main panel. A more fundamental constraint is the recency of the data: Sentinel-5P became operational only in 2018, restricting the observation window to approximately seven years and precluding the longer panels that would support more powerful tests of temporal stability and cross-cycle variation. Finally, the spatial resolution of Sentinel-5P TROPOMI at approximately 3.5×5.5 km per pixel means that each observation integrates atmospheric contributions over a ground area larger than the active pit at most mines, attenuating the mine-specific signal and limiting the precision of the satellite-production relationship relative to what finer-resolution instruments could theoretically deliver.

Future research could extend this framework in several directions. Expanding the mine panel to additional commodities such as coal and nickel, and to a broader set of operators with sufficient mine-level disclosure, would increase statistical power and allow more precise characterisation of the commodity-level heterogeneity documented here. Expanding the InSAR volume dataset as displacement records become available for more mines would allow a more rigorous test of the physical cross-check mechanism. More fundamentally, the next generation of satellite instruments offering finer spatial resolution and higher revisit frequency would directly address the pixel granularity constraint, potentially unlocking the signal-production relationship at mines and in atmospheric environments where it currently falls below the detection threshold. Together these extensions would build toward a richer and more operationally deployable satellite monitoring framework for the mining sector.

References

Baffes, J., & Nagle, P. (2022, December). *Commodity Markets: Evolution, challenges, and policies*. https://doi.org/10.1596/978-1-4648-1911-7

Bai, J., Philippon, T., & Savov, A. (2016). Have financial markets become more informative? *Journal of Financial Economics*, *122*(3), 625–654.

Bartov, E., Faurel, L., & Mohanram, P. S. (2017). Can Twitter help predict Firm-Level earnings and stock returns? *The Accounting Review*, *93*(3), 25–57.

Begenau, J., Farboodi, M., & Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, *97*, 71–87.

Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., de Laat, A., & Wagner, T. (2021). Catalog of NO_x emissions from point sources as derived from the divergence of the NO_2 flux for TROPOMI. *Earth System Science Data*, *13*(6), 2995–3012.

Beirle, S., Borger, C., Jost, A., & Wagner, T. (2023). Improved catalog of NO_x point source emissions (version 2). *Earth System Science Data*, *15*(7), 3051–3073.

Bluhm, R., & Krause, M. (2022). Top lights: Bright cities and their contribution to economic development. *Journal of Development Economics*, *157*, 102880.

Brancaccio, G., Kalouptsidi, M., & Papageorgiou, T. (2020). Geography, transportation, and endogenous trade costs. *Econometrica*, *88*(2), 657–691.

Burke, M., & Lobell, D. B. (2017). Satellite-based assessment of yield variation and its determinants in smallholder african systems. *Proceedings of the National Academy of Sciences*, *114*(9), 2189–2194.

Cao, S., Jiang, W., Wang, J., & Yang, B. (2024). From Man vs. Machine to Man + Machine: The art and AI of stock analyses. *Journal of Financial Economics*, *160*, 103910.

Chen, H., De, P., Hu, Y., & Hwang, B.-H. (2014). Wisdom of Crowds: the value of stock opinions transmitted through social media. *Review of Financial Studies*, *27*(5), 1367–1403.

Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, *108*(21), 8589–8594.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499.

Darvishi Boloorani, A., Soleimani, M., Papi, R., Nasiri, N., Amiri, F., Neysani Samany, N., Huang, K., Gholami, I., & Al-Hemoud, A. (2025). Global map of characterized dust sources using multisource remote sensing data. *Scientific Reports*, *15*, 29805.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., Van Geffen, J., Boersma, K. F., & Veefkind, P. (2018). Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project. *Atmospheric measurement techniques*, *11*(4), 2395–2426.

Dessaint, O., Foucault, T., & Fresard, L. (2024). Does alternative data improve financial forecasting? The horizon effect. *The Journal of Finance*, *79*(3), 2237–2287.

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

Dugast, J., & Foucault, T. (2018). Data abundance and asset price informativeness. *Journal of Financial Economics*, *130*(2), 367–391.

Dugast, J., & Foucault, T. (2024). Equilibrium data mining and data abundance. *The Journal of Finance*, *80*(1), 211–258.

Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210), 1243089.

Escayo, J., Marzan, I., Martí, D., Tornos, F., Farci, A., Schimmel, M., Carbonell, R., & Fernández, J. (2022). Radar interferometry as a monitoring tool for an active mining area using Sentinel-1 C-band data, case study of Riotinto mine. *Remote Sensing*, *14*(13), 3061.

Froot, K., Kang, N., Ozik, G., & Sadka, R. (2017). What do measures of real-time corporate sales say about earnings surprises and post-announcement returns? *Journal of Financial Economics*, *125*(1), 143–162.

Gibson, J., Olivia, S., Boe-Gibson, G., & Li, C. (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics*, *149*, 102602.

Green, T. C., Huang, R., Wen, Q., & Zhou, D. (2019). Crowdsourced employer reviews and stock returns. *Journal of Financial Economics*, *134*(1), 236–251.

Grennan, J., & Michaely, R. (2020). FinTechs and the market for financial analysis. *Journal of Financial and Quantitative Analysis*, *56*(6), 1877–1907.

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *RePEc: Research Papers in Economics*, *70*(3), 393–408.

Hao, X., Wang, Y., Wu, C., & Wu, L. (2024). Oil information uncertainty and aggregate market returns: A natural experiment based on satellite data. *Journal of Financial Markets*, *69*, 100913.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, *102*(2), 994–1028.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.

Katona, Z., Painter, M., Patatoukas, P. N., & Zeng, J. (2025). On the capital market consequences of big data: Evidence from outer space. *Journal of Financial and Quantitative Analysis*, *60*(2), 551–579.

Ma, J., Zhang, R., Li, L., Liu, Z., Sun, J., Kong, L., & Liu, X. (2022). Analysis of the dust-concentration distribution law in an open-pit mine and its influencing factors. *ACS Omega*, *7*(48), 43609–43620.

Martinez-Alonso, S., Veefkind, J. P., Dix, B., Gaubert, B., Theys, N., Granier, C., Soulie, A., Darras, S., Eskes, H., Tang, W., Worden, H., de Gouw, J., & Levelt, P. F. (2023). S-5P/TROPOMI-Derived NOx Emissions From Copper/Cobalt Mining and Other Industrial Activities in the Copperbelt (Democratic Republic of Congo and Zambia). *Geophysical Research Letters*, *50*, e2023GL104109.

Mukherjee, A., Panayotov, G., & Shon, J. (2021). Eye in the sky: Private satellites and government macro data. *Journal of Financial Economics*, *141*(1), 234–254.

Parubets, S., & Naito, H. (2025). Predicting economic activity using atmospheric nitrogen dioxide (NO_2) satellite data: Evidence from local economic indicators in Japan. *PLOS One*, *20*(12), e0337901.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139–1168.

Tetlock, P. C., Saar-tsechansky, M., & Macskassy, S. (2008). More than words: quantifying language to measure firms' fundamentals. *The Journal of Finance*, *63*(3), 1437–1467.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28.

Yang, Y., Zhou, W., Wang, Z., Jiskani, I. M., & Yang, Y. (2024). Accurate long-term dust concentration prediction in open-pit mines: A novel machine learning approach integrating meteorological conditions and mine production intensity. *Journal of Cleaner Production*, *436*, 140411.

Zhu, C. (2019). Big data as a governance mechanism. *Review of Financial Studies*, *32*(5), 2021–2061.

7 Tables and figures

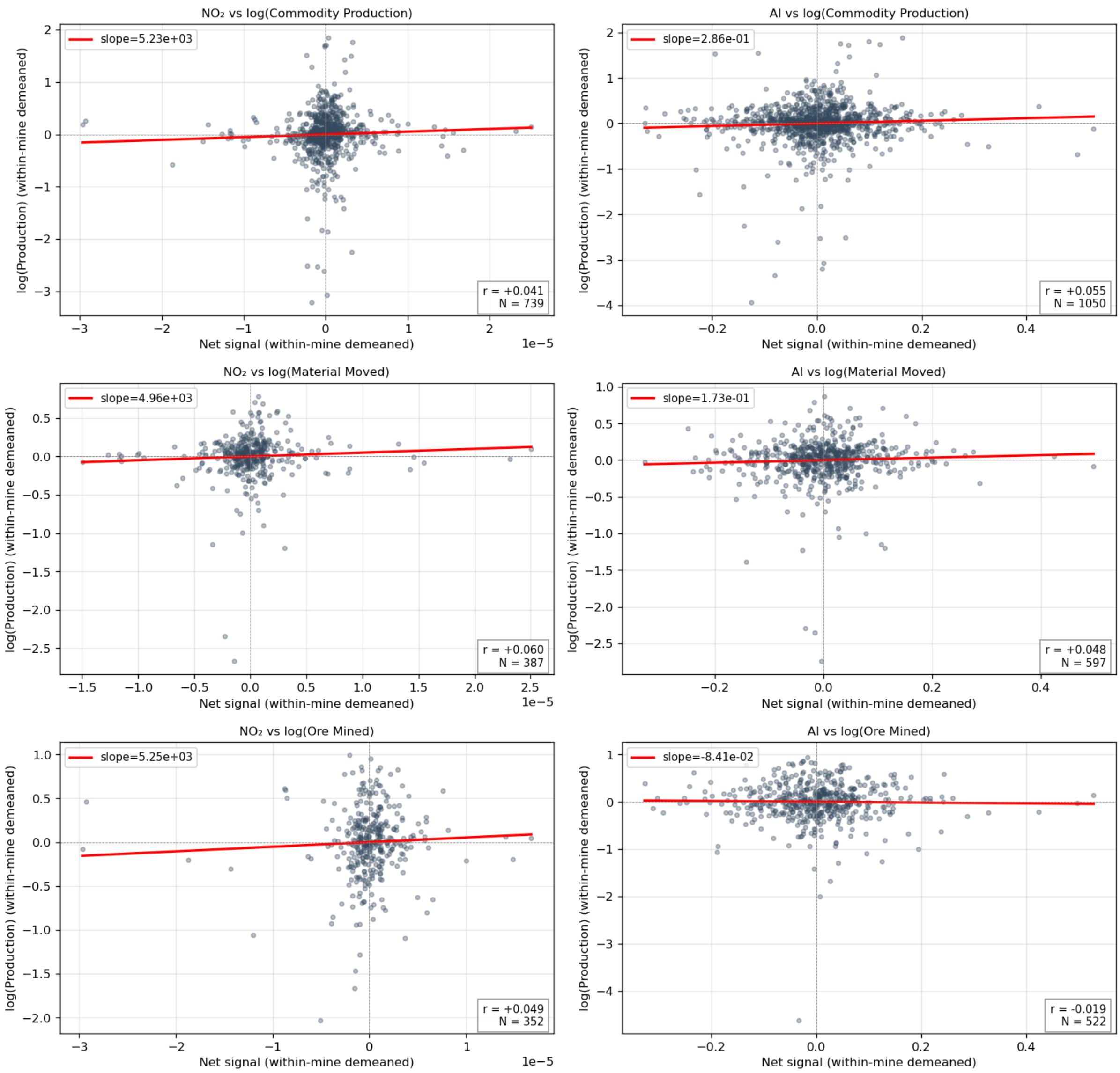

Figure 1: Within-mine demeaned scatter plots of log production against net satellite signal (bbox-control) for all six signal-outcome combinations. Both axes represent deviations from mine-specific means, isolating the within-mine variation exploited by the TWFE and first-differences estimators. The red line shows the pooled OLS fit with slope and within-mine correlation reported in each panel. NO_2 correlations are $r = +0.041$ (*Commodity Production*), $r = +0.060$ (*Material Moved*), and $r = +0.049$ (*Ore Mined*). The aerosol index yields $r = +0.055$, $r = +0.048$, and $r = -0.019$ respectively.

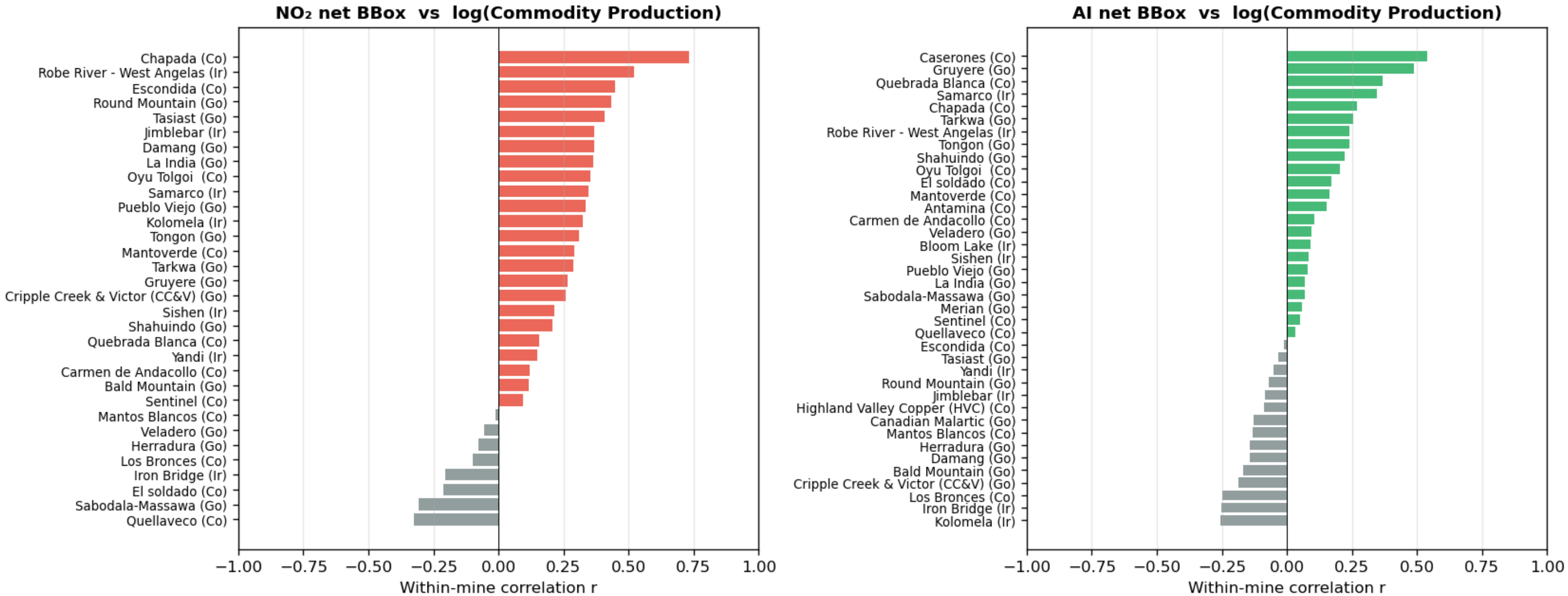

Figure 2: Within-mine time-series correlations between net satellite signal (bbox-control) and log(Commodity Production) for each mine. Left panel: NO_2 net signal (32 mines with ≥ 10 valid quarters). Right panel: aerosol index net signal (38 mines; Lafigué is excluded due to insufficient commodity production observations). Positive bars (coloured) indicate that the satellite signal co-moves with production within that mine over time; negative bars (grey) indicate the opposite. Mines are ranked by correlation within each panel.

Table 1: Contemporaneous TWFE estimates: log(Commodity Production) on Net signal. All specifications include mine fixed effects and quarter fixed effects with mine-clustered standard errors in parentheses. NO_2 specifications include only mine-quarters with at least 30 valid cloud-free days; AI specifications apply the same per-quarter threshold. Zeros in production are treated as missing in the log transformation.

	Panel A: NO_2 net BBox (Spec 1)			*Panel B: Aerosol Index net BBox (Spec 1)*		
	(1) Commodity Production	(2) Material Moved	(3) Ore Mined	(4) Commodity Production	(5) Material Moved	(6) Ore Mined
Net signal	7,036	3,409	7,538	0.303	0.149	−0.079
	(4,631)	(5,250)	(6,040)	(0.207)	(0.135)	(0.157)
Mine FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	739	387	352	1,050	597	522
Mines	38	24	22	39	25	22

	Panel C: Joint specification — NO_2 and AI jointly (Spec 2)		
	Commodity Prod.	Material Moved	Ore Mined
	(7)	(8)	(9)
NO_2 net signal	7,188	3,664	7,996
AI net signal	0.237	0.066	−0.152
N	739	387	352
Mine FE		Yes	
Time FE		Yes	

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Mine-clustered standard errors in parentheses. No coefficients in Panels A and B reach conventional significance thresholds. Panel C reports point estimates only; the joint specification is included to confirm stability of individual signal coefficients when both signals enter simultaneously.

Table 2: First-differences estimates: $\Delta \log$(Commodity Production) on ΔNet signal. Mine fixed effects are absorbed by differencing; quarter fixed effects are included. Mine-clustered standard errors in parentheses. NO_2 specifications include only mine-quarters with at least 30 valid cloud-free days.

	NO_2 net BBox		Aerosol Index net BBox	
	(1) Commodity Production	(2) Ore Mined	(3) Commodity Production	(4) Ore Mined
Δ Net signal	2,690***	5,632	0.134	0.006
	(1,030)	(4,550)	(0.106)	(0.130)
Mine FE (differenced)	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
N	701	330	1,011	500

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Mine-clustered standard errors in parentheses.

Table 3: Heterogeneity in first-differences estimates: $\Delta \log$(Commodity Production) and $\Delta \log$(Ore Mined) on ΔNO_2 net bbox. All specifications include quarter fixed effects with mine-clustered standard errors in parentheses. The pooled estimates reproduce the headline first-differences results from table 2 for reference. Aerosol index results are uniformly insignificant across all heterogeneity dimensions and are omitted for brevity.

	(1) $\Delta \log$(Commodity Prod.)	(2) $\Delta \log$(Ore Mined)
Pooled (reference)		
All mines	2,690***	5,632
	(1,031) [$N = 701$]	(4,554) [$N = 330$]
Panel A: By commodity		
Copper	2,627*	9,908
	(1,387) [$N = 248$]	(6,159) [$N = 154$]
Gold	−508	−6,850
	(3,764) [$N = 291$]	(11,269) [$N = 166$]
Iron ore	10,774***	—
	(4,172) [$N = 162$]	
Panel B: By mine size		
Small mines	1,211	30,657**
	(3,169) [$N = 338$]	(14,974) [$N = 191$]
Large mines	2,115	1,443
	(1,429) [$N = 363$]	(1,224) [$N = 139$]
Panel C: By sub-period		
Pre-2022	5,330**	4,915
	(2,123) [$N = 284$]	(4,307) [$N = 111$]
2022+	1,207	5,235
	(2,092) [$N = 387$]	(9,351) [$N = 203$]
Panel D: By NO_2 data quality		
High retention (≥80%)	3,644**	13,573
	(1,618) [$N = 538$]	(12,462) [$N = 222$]
Medium retention (40–80%)	128	3,594*
	(3,146) [$N = 131$]	(2,056) [$N = 89$]
Time FE	Yes	Yes

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Mine-clustered standard errors in parentheses. Sample sizes in square brackets. Mine size is classified by a within-commodity median split of average quarterly production. Data quality is measured by the share of quarters with at least 30 cloud-free NO_2 observations. Iron ore ore mined is omitted due to insufficient joint observations.

Table 4: Robustness battery: log(Commodity Production) and log(Ore Mined) on Net signal across ten alternative specifications. Coefficients reported with mine-clustered standard errors in parentheses unless otherwise noted. All specifications include mine fixed effects (absorbed in FD) and quarter or commodity-by-quarter fixed effects. Polygon specifications restrict to mines with at least five valid quarters under polygon geometry.

	NO_2 net BBox		Aerosol Index net BBox	
Specification	(1) Commodity Production	(2) Ore Mined	(3) Commodity Production	(4) Ore Mined
Panel A: TWFE level specifications				
(1) Headline TWFE	7,036	7,538	0.303	−0.079
	(4,631)	(6,035)	(0.207)	(0.157)
(2) Polygon geometry (matched)	1,582	2,575	0.235	−0.081
	(3,414)	(5,204)	(0.225)	(0.171)
(3) 50% cloud threshold	7,947	8,932	0.301	−0.081
	(4,937)	(5,883)	(0.207)	(0.156)
(4) Firm-level clustering	7,036	7,538**	0.303	−0.079
	(6,155)	(3,561)	(0.208)	(0.129)
(5) Two-way clustering	7,036	7,538	0.303	−0.079
	(5,006)	(6,439)	(0.205)	(0.139)
(6) Commodity × quarter FE	1,138	9,176**	0.197	−0.070
	(3,945)	(4,425)	(0.183)	(0.171)
Panel B: First-differences specifications				
(7) FD Headline	2,690***	5,632	0.134	0.006
	(1,031)	(4,554)	(0.107)	(0.130)
(8) FD Polygon (matched)	3,659***	7,975	0.120	−0.040
	(742)	(5,842)	(0.129)	(0.153)
(9) FD firm-clustered	2,690***	5,632**	0.134	0.006
	(679)	(2,701)	(0.090)	(0.090)
(10) FD commodity × quarter FE	2,671**	7,799	0.101	0.023
	(1,259)	(5,270)	(0.101)	(0.113)
Sign consistent with headline	10/10	10/10	10/10	7/10
Significant at $p < 0.10$	4/10	3/10	0/10	0/10

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Mine-clustered standard errors in parentheses unless otherwise noted. Specification (4) and (9) use firm-level clustering (17 and 16 firms respectively). Specification (5) uses two-way mine-by-quarter clustering. Specification (6) and (10) replace year-quarter fixed effects with commodity-by-quarter fixed effects. Polygon specifications in (2) and (8) restrict to the 25 mines with at least five valid NO_2 quarters under the polygon geometry.

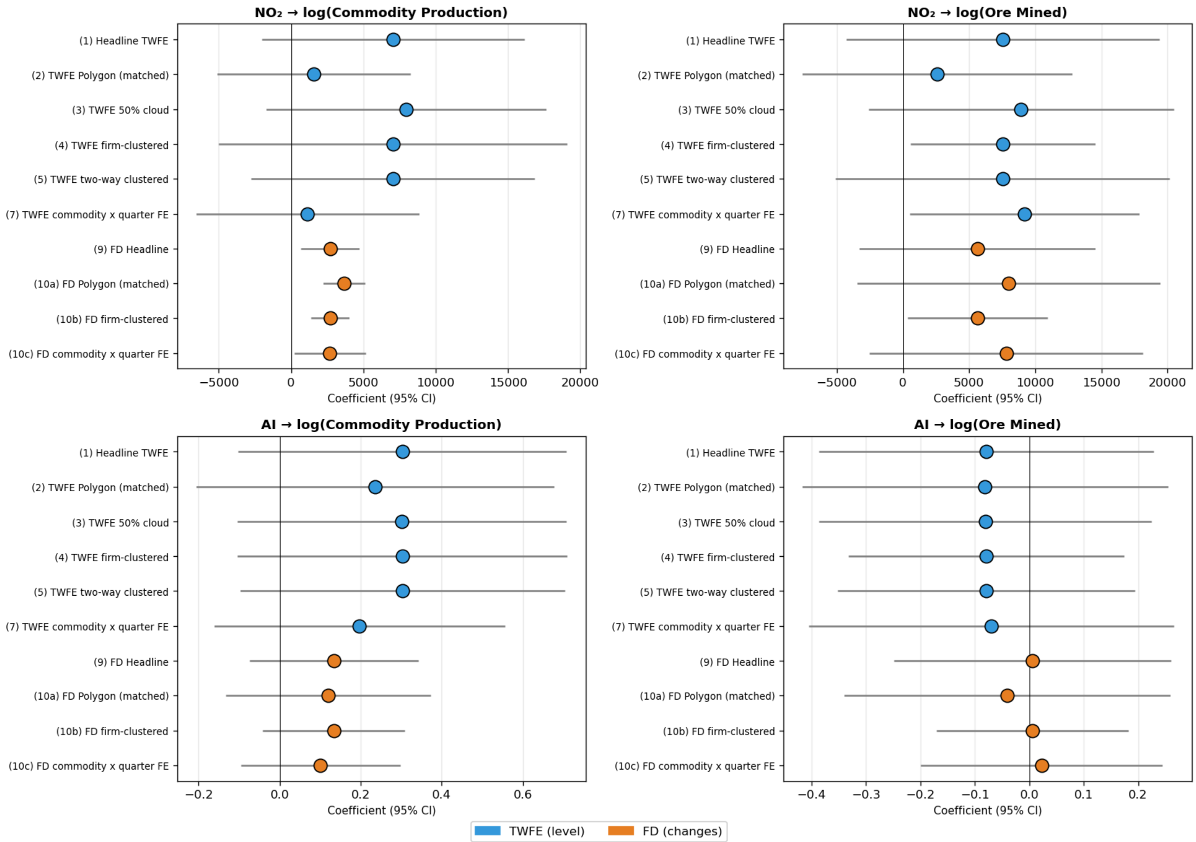

Figure 3: Robustness battery: point estimates with 95 percent confidence intervals across ten specifications for NO_2 and aerosol index signals against log(Commodity Production) and log(Ore Mined). Blue markers denote TWFE level specifications; orange markers denote first-differences specifications. The vertical line at zero serves as a reference. All NO_2 specifications return positive coefficients for both dependent variables. No aerosol index coefficient reaches conventional significance in any specification.

Table 5: Logit estimates of earnings surprise direction on pre-announcement NO_2 satellite signal. The dependent variable is a binary indicator equal to one if adjusted reported EPS exceeds the adjusted Bloomberg consensus estimate. The sample includes firms with at least 50 percent of mines represented in the satellite panel.

	(1) Continuous Signal	(2) Above-Median Signal
Signal	0.392*	0.936**
	(0.213)	(0.394)
Half-yearly reporter	0.062	0.097
	(0.455)	(0.452)
Intercept	−0.047	−0.526
	(0.214)	(0.321)
Odds ratio (signal)	1.48	2.55
N	126	126
Pseudo R^2	0.024	0.036
LLR p-value	0.125	0.042

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors in parentheses. Specification (1) uses the continuous standardised NO_2 signal; specification (2) uses a binary indicator equal to one if the firm's signal exceeds its own historical median.

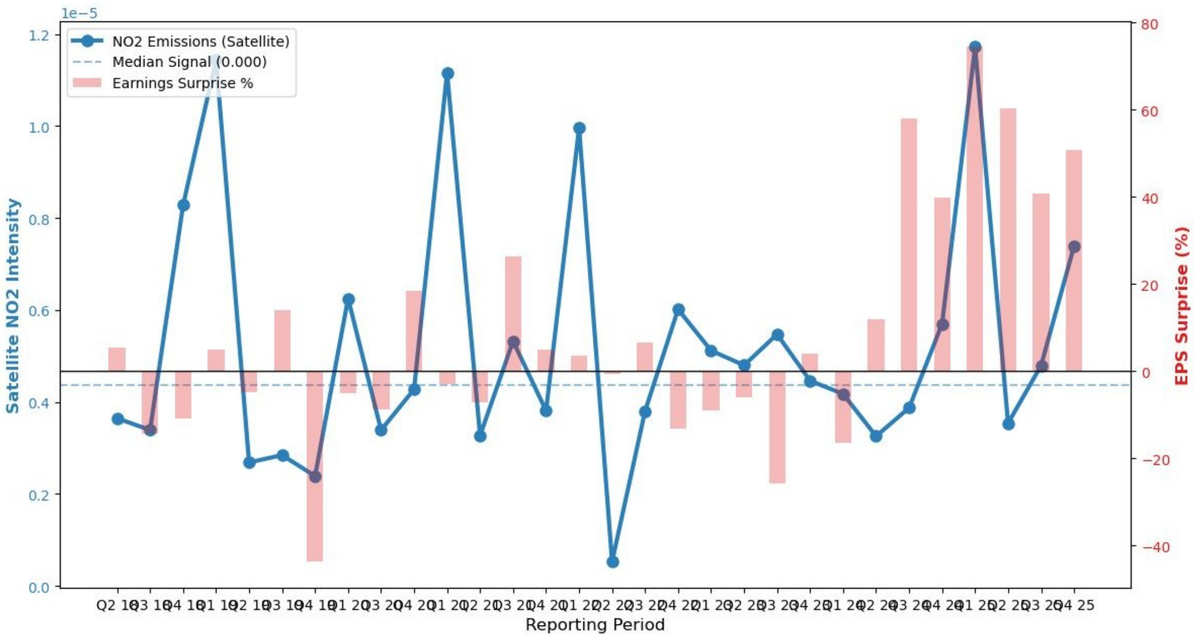

Figure 4: Operational intensity versus earnings surprise for Teck Resources, Q2 2018 to Q4 2025. The blue line plots the quarterly NO_2 net bbox signal averaged across Teck's three sample mines; the dashed line indicates the firm's historical median signal. Red bars report the adjusted EPS surprise percentage at each earnings announcement. The within-firm correlation between the satellite signal and the EPS surprise is $r = +0.52$.

Table 6: Encompassing regression: realised EPS on satellite production forecast and analyst consensus. The dependent variable is adjusted reported EPS. The fitted production forecast is constructed as $\hat{\beta}_{FD} \cdot \Delta NO_2$ from the first-differences model of section 5.4, aggregated to the firm-period level. All specifications include firm fixed effects with standard errors clustered at the firm level in parentheses.

	(1) Forecast only	(2) Consensus only	(3) Encompassing
Panel A: High-coverage firms (6 firms, 5 effective)			
Fitted $\Delta \log(\text{Prod})$	+2.732		+11.759**
	(4.102)		(5.827)
Consensus estimate		+1.011***	+1.015***
		(0.056)	(0.056)
Observations	87	87	87
Within R^2	0.000	0.857	0.862
Firms	5	5	5
Panel B: Full sample (robustness)			
Fitted $\Delta \log(\text{Prod})$	+0.489		+2.261
	(3.777)		(2.607)
Consensus estimate		+1.003***	+1.003***
		(0.034)	(0.034)
Observations	236	236	236
Within R^2	0.000	0.874	0.874
Firms	16	16	16
Firm FE	Yes	Yes	Yes

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Firm-clustered standard errors in parentheses. The fitted production forecast equals $\hat{\beta}_{FD} \cdot \Delta NO_2$, where $\hat{\beta}_{FD}$ is from the first-differences specification in Table 2. For half-yearly reporters the forecast is averaged across the two constituent quarters. A half-year reporter dummy is omitted as it is absorbed by firm fixed effects. Panel A restricts to the six firms with at least 50 percent mine coverage in the satellite panel; Panel B reports the full 17-firm sample as a robustness check.

Table 7: Trading strategy performance: long-short returns based on pre-announcement NO_2 signal for high-coverage firms. The sample covers six firms with at least 50 percent mine coverage. Variant A uses the full-sample within-firm median as threshold (in-sample). Variant B uses an expanding within-firm median (out-of-sample).

	Variant A (In-sample)	Variant B (Out-of-sample)
Mean per-trade return	−0.224%	−0.045%
t-statistic	−0.52	−0.10
p-value	0.606	0.924
Sharpe ratio	−0.049	−0.010
Hit rate	47.3%	49.0%
Long-leg mean return	+0.336%	+0.586%
Short-leg mean return	+0.784%	+0.650%
Long minus short	−0.448%	−0.064%
Bootstrap 95% CI	—	[−1.420%, +1.036%]
Bootstrap p-value	—	0.967
N trades	110	100

Notes: Firms in sample: Fortescue, Anglo American, Kinross Gold, Teck Resources, Capstone Copper, Champion Iron. Variant B first tradeable date: 23 April 2019 after ten-observation burn-in period. Bootstrap inference uses firm-block resampling with 5,000 replications.

Table 8: 2x2 decomposition of announcement-day returns by signal class and earnings outcome (Variant B, out-of-sample). Mean announcement-day returns reported in percent. Beat defined as adjusted reported EPS exceeding adjusted Bloomberg consensus estimate.

	Miss	Beat	P(beat)
High signal	−1.334% ($N = 15$)	+1.433% ($N = 34$)	69.4% ($N = 49$)
Low signal	+0.184% ($N = 31$)	+1.374% ($N = 20$)	39.2% ($N = 51$)
E(long leg)		+0.586%	
E(short leg)		+0.650%	
Long minus short		−0.064%	

Notes: High signal defined as firm NO_2 signal exceeding its own expanding historical median at announcement date. Expected long and short leg returns computed as probability-weighted averages across beat and miss outcomes within each signal class.

Table 9: Pre-announcement trading strategy: contrarian direction. The dependent variable is the mining-adjusted cumulative return over each pre-announcement window. The strategy takes a long position when $\hat{P}(\text{Beat} = 1) < 0.5 - \tau$ and a short position when $\hat{P}(\text{Beat} = 1) > 0.5 + \tau$, where $\hat{P}(\text{Beat} = 1)$ is generated by the expanding-window OOS logit of section 4.5.2. All inference uses firm-block bootstrap with 5,000 replications.

Window	τ	Mean	Boot p	CI low	CI high	Sharpe	Hit %	N	Long / Short
Panel A: $[-1, +1]$	0.00	+1.275%*	0.051	−0.01%	+2.70%	+0.215	60.0%	100	67 / 33
	0.05	+1.722%**	0.039	+0.05%	+3.23%	+0.289	61.8%	68	45 / 23
	0.10	+1.635%	0.206	−0.94%	+3.88%	+0.248	65.9%	44	32 / 12
	0.15	+1.744%	0.318	−1.82%	+4.54%	+0.230	66.7%	27	21 / 6
Panel B: $[-5, +1]$	0.00	+0.743%	0.324	−0.74%	+2.56%	+0.115	53.0%	100	67 / 33
	0.05	+1.318%	0.247	−0.92%	+3.24%	+0.195	51.5%	68	45 / 23
	0.10	+1.640%	0.293	−1.99%	+4.36%	+0.209	54.5%	44	32 / 12
	0.15	+2.049%	0.359	−2.95%	+5.22%	+0.226	59.3%	27	21 / 6
Panel C: $[-10, +1]$	0.00	+0.142%	0.791	−1.69%	+1.82%	+0.019	48.0%	100	67 / 33
	0.05	+1.089%	0.400	−1.58%	+3.10%	+0.142	51.5%	68	45 / 23
	0.10	+1.089%	0.438	−2.43%	+3.39%	+0.125	50.0%	44	32 / 12
	0.15	+1.936%	0.399	−3.62%	+5.55%	+0.191	59.3%	27	21 / 6

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Firm-clustered bootstrap standard errors with 5,000 replications. Mean returns are mining-adjusted (excess over iShares MSCI Global Metals and Mining Producers ETF). The strategy goes long when $\hat{P}(\text{Beat} = 1) < 0.5 - \tau$ and short when $\hat{P}(\text{Beat} = 1) > 0.5 + \tau$, fading the direction of the satellite signal; observations with $|\hat{P}(\text{Beat} = 1) - 0.5| \leq \tau$ are excluded. Sample: six high-coverage firms, expanding OOS logit with burn-in of 20 announcements. Long and short counts reflect the number of contrarian long and short positions respectively.

Figure 5: Cumulative long-short strategy return over the out-of-sample period. Each panel plots the simple cumulative sum of mining-adjusted per-trade returns at $\tau = 0.00$ for the three pre-announcement windows $[-1, +1]$, $[-5, +1]$, and $[-10, +1]$. The strategy takes a long position when $\hat{P}(\text{Beat} = 1) < 0.5 - \tau$ and a short position when $\hat{P}(\text{Beat} = 1) > 0.5 + \tau$, based on the expanding-window OOS logit. The shaded area reflects the sign of the cumulative return at each point in time. Final cumulative returns are +127.5 percent, +74.3 percent, and +14.2 percent for the $[-1, +1]$, $[-5, +1]$, and $[-10, +1]$ windows respectively.

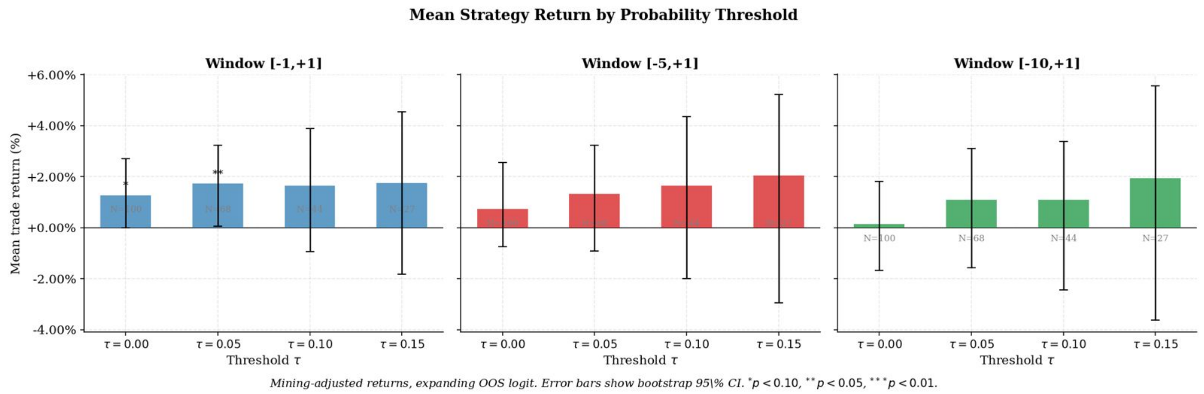

Figure 6: Mean per-trade mining-adjusted return by probability threshold τ across three pre-announcement windows. Bars report the mean trade return at each threshold level; error bars show bootstrap 95 percent confidence intervals based on firm-block resampling with 5,000 replications. The number of trades N is reported inside each bar. All returns are positive across threshold levels in the $[-1, +1]$ window, with the $\tau = 0.05$ specification significant at the five percent level. Returns in the $[-5, +1]$ and $[-10, +1]$ windows are positive but uniformly insignificant. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix A Data and sample

Table 10: Mining operations by company and location. The sample covers 39 open-pit mines operated by 17 corporations across 17 countries, spanning the copper, gold, and iron ore sectors. All mines are observed over the satellite panel period Q2 2018 to Q4 2025. Mine selection follows the criteria described in section 3.2.

Company	**Mine**	**Country**	**Commodity**
Agnico Eagle	La India	Mexico	Gold
Agnico Eagle	Canadian Malartic	Canada	Gold
Anglo American	Los Bronces	Chile	Copper
Anglo American	El Soldado	Chile	Copper
Anglo American	Quellaveco	Peru	Copper
Anglo American	Sishen	South Africa	Iron ore
Anglo American	Kolomela	South Africa	Iron ore
Barrick Gold	Tongon	Côte d'Ivoire	Gold
Barrick Gold	Pueblo Viejo	Dominican Rep.	Gold
Barrick Gold	Veladero	Argentina	Gold
BHP Group	Escondida	Chile	Copper
BHP Group	Antamina	Peru	Copper
BHP Group	Samarco	Brazil	Iron ore
BHP Group	Yandi	Australia	Iron ore
BHP Group	Jimblebar	Australia	Iron ore
Capstone Copper	Mantos Blancos	Chile	Copper
Capstone Copper	Mantoverde	Chile	Copper
Champion Iron	Bloom Lake	Canada	Iron ore
Endeavour Mining	Sabodala-Massawa	Senegal	Gold
Endeavour Mining	Lafigué	Côte d'Ivoire	Gold
First Quantum	Sentinel	Zambia	Copper
Fortescue	Iron Bridge	Australia	Iron ore
Fresnillo	Herradura	Mexico	Gold
Gold Fields	Tarkwa	Ghana	Gold
Gold Fields	Damang	Ghana	Gold
Gold Fields	Gruyere	Australia	Gold
Kinross Gold	Tasiast	Mauritania	Gold
Kinross Gold	Round Mountain	USA	Gold
Kinross Gold	Bald Mountain	USA	Gold
Lundin Mining	Caserones	Chile	Copper
Lundin Mining	Chapada	Brazil	Copper
Newmont	Cripple Creek	USA	Gold
Newmont	Merian	Suriname	Gold
Pan American	Shahuindo	Peru	Gold
Rio Tinto	Robe River	Australia	Iron ore
Rio Tinto	Oyu Tolgoi	Mongolia	Copper
Teck Resources	Highland Valley	Canada	Copper
Teck Resources	Quebrada Blanca	Chile	Copper
Teck Resources	Carmen de Andacollo	Chile	Copper

Table 11: Production summary statistics by commodity, Q2 2018 to Q4 2025. Observations are at the mine-quarter level and restricted to non-zero positive values. *Material Moved* and *Ore Mined* are reported in tonnes. *Commodity Production* is reported in tonnes of contained metal for copper, tonnes of concentrate for iron ore, and troy ounces for gold.

Metric	**Commodity**	**N**	**Mines**	**Mean**	**Std Dev**	**P25**	**Median**	**P75**
Material Moved	Copper	218	9	35.47M	33.04M	9.66M	20.42M	56.83M
	Gold	337	14	13.82M	6.87M	8.97M	11.78M	17.57M
	Iron ore	42	2	12.01M	4.93M	8.62M	10.40M	15.68M
	All	597	25	21.59M	23.17M	9.23M	13.74M	22.42M
Ore Mined	Copper	196	9	6.63M	4.84M	1.91M	7.07M	10.69M
	Gold	284	11	3.15M	1.78M	1.75M	2.93M	4.05M
	Iron ore	42	2	6.67M	2.71M	4.92M	5.64M	9.38M
	All	522	22	4.74M	3.75M	1.88M	3.51M	6.69M
Commodity Production	Copper	356	14	60.5k	76.3k	12.6k	33.7k	71.0k
	Gold	477	17	88.5k	62.2k	45.7k	73.4k	123.6k
	Iron ore	217	8	7.26M	5.74M	2.87M	6.23M	8.48M
	All†	1,050	39	1.56M	3.90M	36.3k	77.8k	186.4k

Notes: "M" denotes millions, "k" denotes thousands. †The "All" row for Commodity Production aggregates across commodities with different reporting units; cross-commodity comparison of mean magnitudes is not meaningful and is reported for completeness only.

Table 12: Satellite signal summary statistics under the primary bbox-control geometry. Raw signals measure the mine-area atmospheric column; net signals subtract the corresponding daily control-area measurement to isolate mine-specific atmospheric activity. Sample restricted to mine-quarters with at least 30 valid cloud-free days.

Variable	**N**	**Mean**	**Std Dev**	**P25**	**Median**	**P75**
NO_2 raw BBox (mol/m^2)	859	$+1.85 \times 10^{-5}$	$+1.17 \times 10^{-5}$	$+1.13 \times 10^{-5}$	$+1.47 \times 10^{-5}$	$+2.18 \times 10^{-5}$
NO_2 net BBox (mol/m^2)	859	$+4.15 \times 10^{-6}$	$+7.31 \times 10^{-6}$	$+1.19 \times 10^{-6}$	$+2.83 \times 10^{-6}$	$+5.84 \times 10^{-6}$
AI raw BBox (unitless)	1,209	+0.0294	+0.5867	−0.3655	−0.0909	+0.3419
AI net BBox (unitless)	1,209	+0.1077	+0.2178	−0.0201	+0.0569	+0.1645

Notes: Net signals are computed as the daily difference between mine-area and control-area measurements, then aggregated to quarterly means. NO_2 sample excludes mines with fewer than 10 valid quarters across the panel, yielding 34 mines and 859 mine-quarter observations.

Table 13: Satellite signal summary statistics under the alternative polygon geometry. Net signals subtract the corresponding daily control-area measurement from the mine-area polygon measurement. Sample restricted to mine-quarters with at least 30 valid cloud-free days. Under the polygon geometry, 10 of 39 mines lose all NO_2 observations and 10 mines lose all AI observations due to the mechanical sensitivity of small measurement areas to cloud presence, as discussed in section 3.3.

Variable	**N**	**Mean**	**Std Dev**	**P25**	**Median**	**P75**
NO_2 net Polygon (mol/m^2)	602	$+4.87 \times 10^{-6}$	$+8.39 \times 10^{-6}$	$+1.41 \times 10^{-6}$	$+3.47 \times 10^{-6}$	$+7.11 \times 10^{-6}$
AI net Polygon (unitless)	899	+0.1422	+0.2383	+0.0024	+0.0937	+0.1981

Notes: Net signals are computed as the daily difference between mine-area Polygon and control-area measurements, then aggregated to quarterly means.

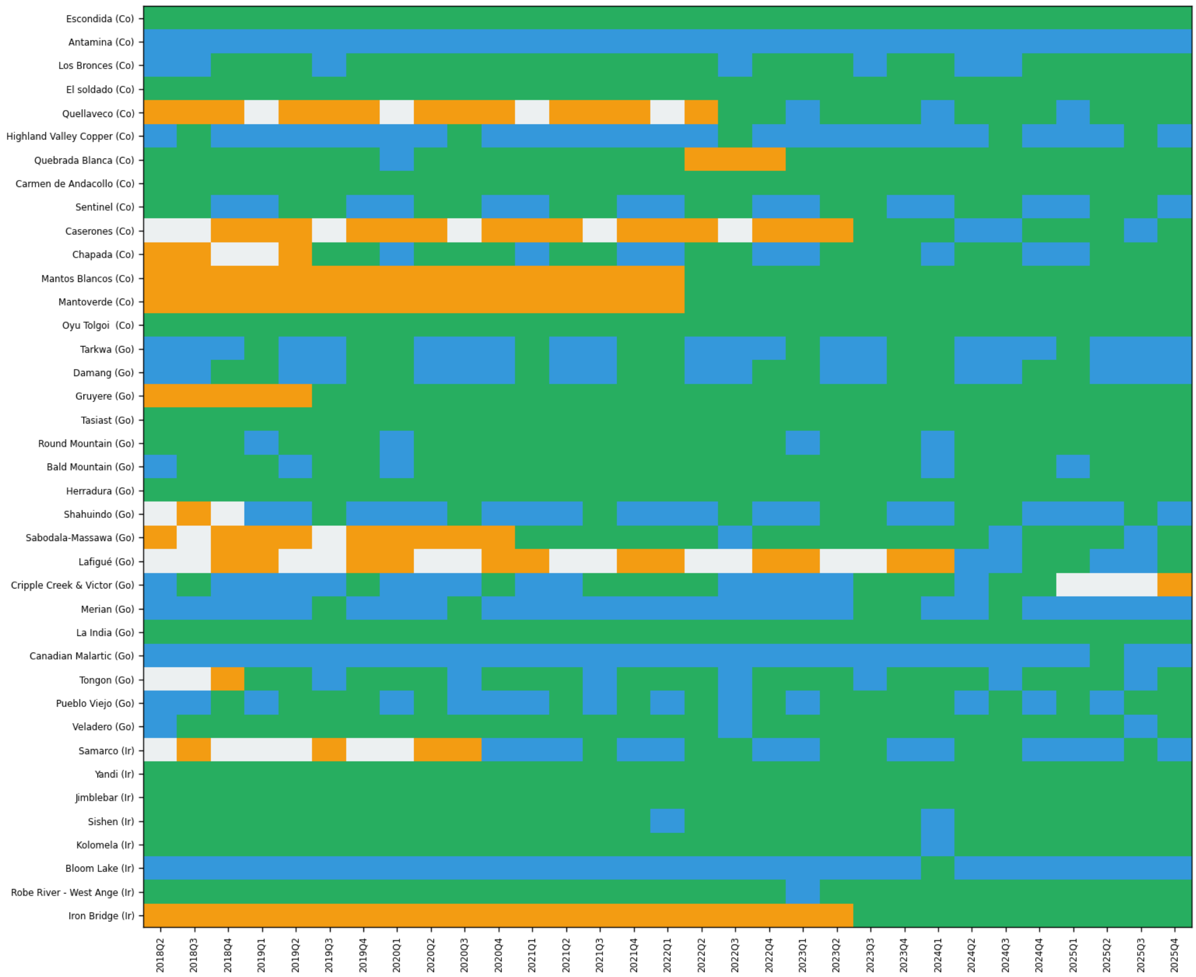

Figure 7: Data coverage by mine and quarter, Q2 2018 to Q4 2025. Green cells indicate quarters where both a valid satellite signal and production data are available; blue cells indicate valid satellite signal but missing production; orange cells indicate available production but insufficient cloud-free days; white cells indicate neither observation is available. Mine labels denote commodity type: Co = Copper, Go = Gold, Ir = Iron ore. Persistent orange or white patterns at specific mines reflect chronic cloud cover, operational disruptions, or incomplete production disclosure rather than satellite data failures.

Appendix B Auxiliary estimation results

Table 14: Lead-lag structure: log commodity production on net satellite signal at horizons $t-2$ through $t+2$. Both specifications include mine and quarter fixed effects. NO_2 specification includes only mine-quarters with at least 30 valid cloud-free days.

	NO_2 net BBox		Aerosol Index net BBox	
Horizon	$\hat{\beta}$	SE	$\hat{\beta}$	SE
$t-2$	1,373	(3,290)	−0.018	(0.190)
$t-1$	709	(3,890)	0.125	(0.172)
t	4,099	(3,240)	0.400	(0.309)
$t+1$	2,063	(4,110)	0.087	(0.177)
$t+2$	3,340	(2,650)	0.133	(0.205)
N	648		920	
Mines	37		39	
Mine FE	Yes		Yes	
Time FE	Yes		Yes	

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Mine-clustered standard errors in parentheses. No coefficient at any horizon reaches conventional significance for either signal. The dependent variable is log commodity production in all specifications.

Table 15: Pre-announcement trading strategy: intuitive direction. The dependent variable is the mining-adjusted cumulative return over each pre-announcement window. The strategy takes a long position when $\hat{P}(\text{Beat} = 1) > 0.5 + \tau$ and a short position when $\hat{P}(\text{Beat} = 1) < 0.5 - \tau$, where $\hat{P}(\text{Beat} = 1)$ is generated by the expanding-window OOS logit of section 4.5.2. All inference uses firm-block bootstrap with 5,000 replications.

Window	τ	Mean	Boot p	CI low	CI high	Sharpe	Hit %	N	Long / Short
Panel A: $[-1, +1]$	0.00	−1.275%*	0.051	−2.70%	+0.01%	−0.215	40.0%	100	33 / 67
	0.05	−1.722%**	0.039	−3.23%	−0.05%	−0.289	38.2%	68	23 / 45
	0.10	−1.635%	0.206	−3.88%	+0.94%	−0.248	34.1%	44	12 / 32
	0.15	−1.744%	0.318	−4.54%	+1.82%	−0.230	33.3%	27	6 / 21
Panel B: $[-5, +1]$	0.00	−0.743%	0.334	−2.55%	+0.74%	−0.115	47.0%	100	33 / 67
	0.05	−1.318%	0.230	−3.25%	+0.84%	−0.195	48.5%	68	23 / 45
	0.10	−1.640%	0.278	−4.36%	+1.65%	−0.209	45.5%	44	12 / 32
	0.15	−2.049%	0.344	−5.25%	+3.09%	−0.226	40.7%	27	6 / 21
Panel C: $[-10, +1]$	0.00	−0.142%	0.812	−1.88%	+1.69%	−0.019	52.0%	100	33 / 67
	0.05	−1.089%	0.401	−3.10%	+1.58%	−0.142	48.5%	68	23 / 45
	0.10	−1.089%	0.458	−3.39%	+2.73%	−0.125	50.0%	44	12 / 32
	0.15	−1.936%	0.378	−5.55%	+3.52%	−0.191	40.7%	27	6 / 21

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Firm-clustered bootstrap standard errors with 5,000 replications. Mean returns are mining-adjusted (excess over iShares MSCI Global Metals and Mining Producers ETF). Results are consistent in sign and magnitude under raw and S&P 500-adjusted benchmarks. The strategy goes long when $\hat{P}(\text{Beat} = 1) > 0.5 + \tau$ and short when $\hat{P}(\text{Beat} = 1) < 0.5 - \tau$; observations with $|\hat{P}(\text{Beat} = 1) - 0.5| \leq \tau$ are excluded. Sample: six high-coverage firms, expanding OOS logit with burn-in of 20 announcements.

Appendix C InSAR volume estimates

Table 16: InSAR volume estimates: effect of InSAR-derived pit volume on log production. All specifications include mine and time fixed effects and are estimated with NO_2 and AI as controls. Regressors are standardised to unit variance. Bootstrap p-values use wild cluster bootstrap with Rademacher weights and 5,000 replications

Specification	$\hat{\beta}$ (Volume$_t$)	SE	p_{asym}	p_{boot}
Panel A: log(*Material Moved*)				
MM-Base	+0.067	0.098	0.493	0.537
MM-Joint	+0.069	0.091	0.450	0.570
MM-Lag	+0.079	0.098	0.419	0.525
Volume$_{t-1}$	+0.068	0.076	0.376	0.430
MM-Gold	+0.057	0.085	0.505	0.594
Panel B: log(*Commodity Production*)				
CP-Joint	+0.037	0.007	0.000	0.016**
CP-Lag	+0.033	0.012	0.008	0.061*
Volume$_{t-1}$	+0.007	0.052	0.891	0.840
CP-Gold	+0.039	0.014	0.004	0.118
Mine FE		Yes		
Time FE		Yes		
Clusters		6 (MM-Gold and CP-Gold: 5)		

Notes: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ based on p_{boot} only. NO_2 and AI enter as controls in all specifications except MM-Base; coefficients omitted for brevity. MM-Gold and CP-Gold drop Antamina as a robustness check. Asymptotic p-values are reported alongside bootstrap p-values to illustrate the unreliability of asymptotic inference with six clusters.

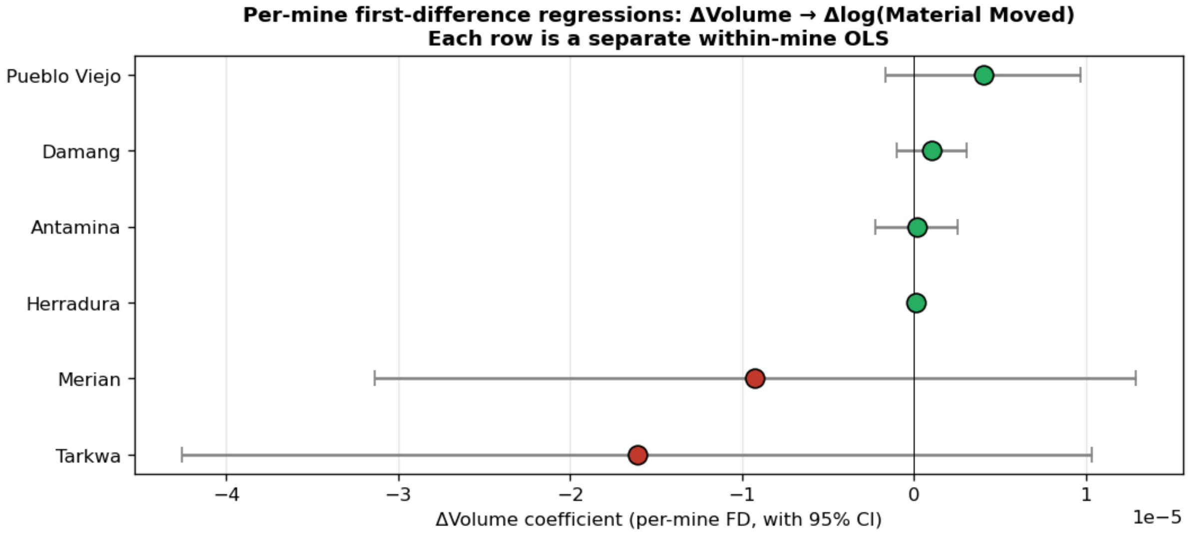

Figure 8: Per-mine first-difference regressions of $\Delta \log(\text{Material Moved})$ on ΔVolume. Each row reports the OLS coefficient and 95 percent confidence interval from a separate within-mine regression. Green markers indicate positive coefficients and red markers indicate negative coefficients. Four of the six mines yield positive estimates, with Pueblo Viejo and Damang showing the strongest positive relationships. Wide confidence intervals throughout reflect the limited within-mine time series available for each site.