

Online Resource 1

Carry Incentives and Net Warrant Accumulation: Evidence from SHFE Aluminum

Yuchen Han^{1*}

^{1*}City University of Macau, Macau, China.

Corresponding author(s). E-mail(s): F24092100015@cityu.edu.mo;

Abstract

Supplementary analyses for SN Business and Economics. This resource reports input reconciliation and historical calendar details; all specified control, liquidity, storage and sample sensitivities; local and crossing-event diagnostics; auxiliary weekly evidence; and reverse predictive regressions. It complements the daily net-warrant analysis in the article and preserves its interpretation limits.

This supplement accompanies the Shanghai Futures Exchange (SHFE) aluminum analysis. Monetary amounts are in Chinese yuan (CNY); inventory and net warrant changes are in tonnes. SE denotes a standard error and FE denotes fixed effects. CRV1 refers to the finite-sample-adjusted cluster-robust covariance estimator. Reported bootstrap p-values use two-sided null-imposed Webb six-point weights and 9,999 draws.

A Input reconciliation and historical calendar

Two weekly observations in February 2025 previously used alumina rather than aluminum because a substring filter admitted both products. The corrected aluminum records are 198,187 tonnes of visible stocks and 86,286 tonnes of warrants on 14 February, and 232,903 and 115,657 tonnes on 21 February. Each warrant figure matches that date's daily aluminum total. On 10 May 2019 the original weekly and daily sources still differ by one tonne, 353,553 versus 353,554; neither source is overwritten to force exact agreement. A sensitivity comparison excludes observations whose outcome or lag depends on that level.

The two recovered daily warrant records are 172,858 tonnes on 2 July 2015 and 404,262 tonnes on 22 April 2019. With these records included, reconstructed first differences agree with the reported daily changes on all usable sessions. The curve is rebuilt from contract settlements, rather than inferred from a saved carry series. The reference contract is excluded from the curve regression, preserving its original contract-set definition.

The historical calendar has three exceptions to the 15th-or-next-session rule within the sample. Original market files confirm the February 2015, 2018, and 2021 contracts on the

selected final date and their absence on the next session. The other 121 completed contract months agree with the general rule in the cached series. The May 2025 nearby contract is right censored by the sample endpoint and uses the ordinary 15 May date. These checks do not replace a complete archive of formal historical adjustment notices.

The legacy foreign-price control is excluded from the revised analysis. It used Yahoo Finance’s CME aluminum futures ticker ALI=F, although its project field was labeled LME, and applied a constant 1.13 gross-up across the sample. That series is not an official LME cash quote or a historically verified import-parity calculation. Reconstructing a cross-border control would require correct instrument identity, synchronized pricing, changing tax treatment, and a defensible transaction-cost definition. No cross-border mechanism claim relies on the legacy proxy here.

B Sensitivity specifications and parameter scenarios

Table S1 reports all defined control, liquidity, sample, and version comparisons. The generic-calendar rows permit comparison with the earlier approximation while the main analysis uses the historical exceptions. The legacy incomplete-sample row reproduces the previously reported coefficient without endorsing its nonconsecutive dates. The alternatives retain their prescribed roles even when a p-value moves across a conventional cutoff.

Table S1 Defined sensitivity and version comparisons

Specification	Difference	SE	Webb p	N
Year FE only	1,517	557	0.035	2,505
Delivery-calendar FE only	1,172	416	0.004	2,505
Exclude 2017–2018	1,950	511	0.024	2,018
Exclude one-tonne discrepancy dependencies	1,368	429	0.007	2,502
Volume-weighted curve	2,045	921	0.019	2,505
Traded 1–6 month curve	2,014	745	0.007	2,505
Generic calendar FE only	1,225	375	0.004	2,505
Generic year and calendar FE	1,244	466	0.096	2,505
Legacy incomplete daily sample	1,380	433	0.008	2,503

Notes. Outcome: daily net warrant change, tonnes. The legacy sample comparison reconstructs the earlier 2,504-day sequence, including its two gaps, solely to document numerical agreement. Generic calendars omit the three Spring Festival overrides. These variants are sensitivity comparisons, not alternate preferred models selected by significance. Missing alternate slopes are excluded before state classification. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

Storage scenarios alter both the mean benchmark and the days classified above it (Table S2). Dated fee records make the scenario range transparent, but charges for different storage and tax arrangements are not interchangeable tariffs for the baseline population. All scenarios retain the same public financing benchmark and omitted-cost limitations.

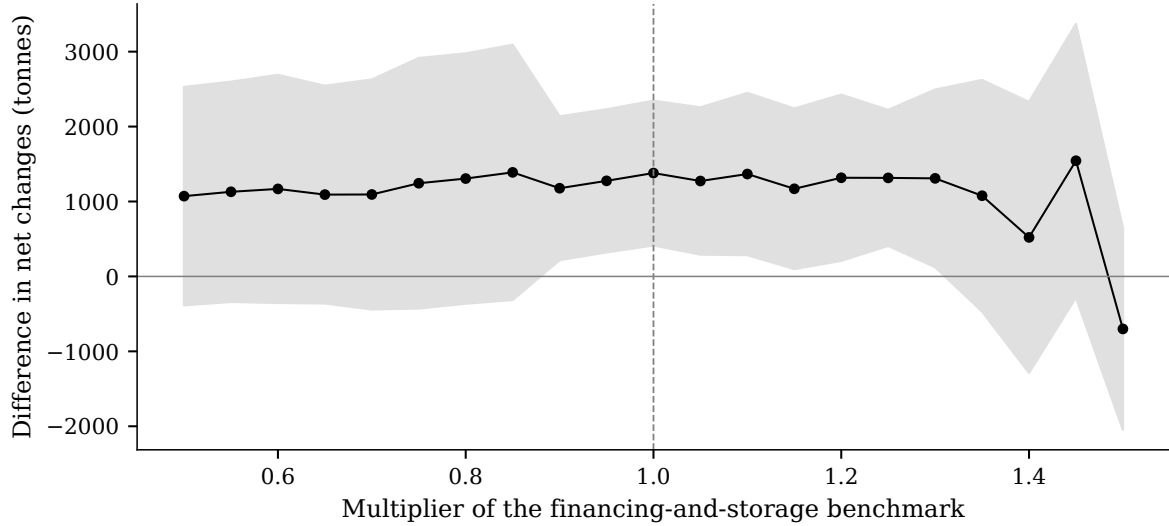
Figure S1 retains the earlier fixed multiplier grid from 0.50 to 1.50 in 0.05 steps. Every point changes the state definition, so the grid is a set of related sensitivity comparisons. The plotted intervals are pointwise approximations and have no simultaneous-coverage

Table S2 Storage-parameter scenarios

s	Mean benchmark	State days	Difference	Webb p
0.40	47.3	324	1,293	0.007
0.45	48.8	300	1,380	0.008
0.60	53.3	230	1,440	0.020

Notes. s is CNY/tonne/day; the benchmark is CNY/tonne/month. Differences are tonnes in an unadjusted lagged state comparison. Each scenario has 2,505 observations and 11 year clusters. The 0.40 and 0.60 values bracket the maintained 0.45 assumption; they are not verified common tariffs for every date in 2015–2025. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p -values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

interpretation. The grid does not select an optimal multiplier or identify one times the benchmark as a special threshold.

**Fig. S1** State comparisons across the fixed benchmark multiplier grid

Notes. Points are coefficients on $\mathbf{1}\{b_{t-1} > qc_{t-1}\}$ for the displayed multiplier q , with outcome ΔW_t in tonnes. Shading is a pointwise 95 percent $t(10)$ approximation interval based on CRV1 year-clustered standard errors. It is not a simultaneous band or a Webb-inverted interval. Every regression has 2,505 observations and eleven year clusters.

Table S3 records each individual year exclusion. These comparisons describe how pooled estimates depend on their time support; they are not independent new samples or a remedy for the concentration of above-benchmark sessions.

Table S3 Leave-one-year-out state comparisons

Excluded year	Difference	SE	Webb p	N
2015	1,514	491	0.020	2,262
2016	1,294	456	0.020	2,261
2017	1,338	407	0.029	2,261
2018	1,990	394	0.007	2,262
2019	1,231	457	0.013	2,261
2020	1,198	378	0.022	2,262
2021	1,431	423	0.010	2,262
2022	1,298	437	0.016	2,263
2023	1,402	441	0.011	2,263
2024	1,324	438	0.018	2,263
2025	1,388	436	0.009	2,430

Notes. Unadjusted differences in next-session net warrant changes, tonnes. Each row excludes one year and retains ten year clusters. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

C Local and event diagnostic details

The level shift and change in slope in equation (6) in the article are distinct hypotheses (Table S4). The windows are fixed before the revised estimates, and no alternative bandwidth is substituted because of its significance. The comparison requires neither gross-registration data nor a causal threshold assumption, and establishes neither.

Table S4 Jump and change-in-slope diagnostics

Window	Jump (tonnes)	Webb p	Slope change	Webb p
Full sample	1,479	0.080	-5.98	0.646
± 10	1,696	0.091	125.64	0.687
± 20	1,006	0.419	-24.22	0.731
± 40	705	0.479	-36.00	0.195

Notes. Equation (6) includes a lagged continuous gap, the lagged state indicator, and their interaction. The slope change is tonnes per CNY/tonne/month. Samples and year counts are reported in Table 3 in the article. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

Table S5 shows the corrected daily dates in the 2018 illustration. A spell's signal interval and its response interval differ by one trading session: a signal from t_0 through t_1 corresponds to responses from $t_0 + 1$ through $t_1 + 1$. Cumulative changes are computed on those response dates. This prevents an event description from mixing date conventions.

Table S5 Corrected dates in the June–July 2018 illustration

Response date	Nearby contract	Calendar days to last	Net change
2018-06-15	AL1806	0	5,886
2018-07-13	AL1807	3	4,794
2018-07-16	AL1807	0	-515

Notes. Net changes are tonnes. July 13 is three calendar days before the last trading date; July 16 is the last date. Warrant cancellations and gross registrations are not observed separately, so these rows do not identify physical metal movements or delivery-induced outflows.

D Auxiliary weekly evidence

Figure S2 displays the corrected weekly stock series. These variables come from the same reports, while the daily warrant totals and daily fitted slopes supply the main analysis. Neither a frequency change nor aggregation produces an independent dataset.

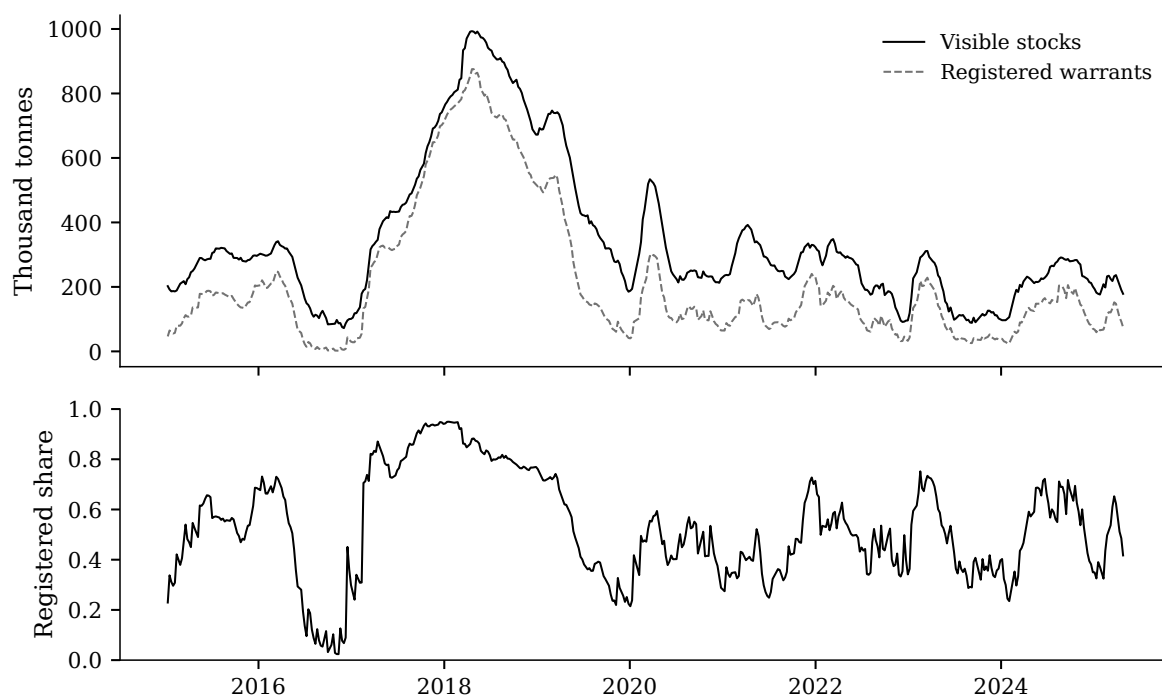


Fig. S2 Visible stocks, registered warrants, and their share

Notes. Original weekly aluminum grand totals, 9 January 2015–25 April 2025, 496 reports. The upper panel is thousand tonnes and the lower panel is W/S . Aluminum is selected by exact product identity, excluding alumina. Source: SHFE weekly warehouse reports.

Table S6 relates the same-Friday slope to registered warrants, controlling for visible stocks and scaling both stock measures by 100,000 tonnes. The corrected coefficient on W is 20.01, with Webb p -value 0.045. The legacy row documents the consequence of the two incorrectly selected weekly totals. This levels association is auxiliary: it can reflect persistent common conditions and does not verify the interpretation of the daily net-change model. A next-session contract-panel coefficient can differ from this collapsed Friday regression because timing, contract support, and weighting differ.

Table S6 Auxiliary weekly stock-composition association

Weekly input	Coefficient on W	SE	Webb p	N
Corrected exact-AL weekly totals	20.01	9.40	0.045	496
Legacy substring selection	19.99	9.31	0.041	496

Notes. The outcome is the same-Friday curve slope. Both regressions control for weekly visible stocks S and scale S, W by 100,000 tonnes. The corrected input uses only aluminum ($\text{VARID} = al$, or an exact aluminum name in older reports); the legacy version incorrectly selected alumina in two weeks. A levels association is separate from the daily net-change estimand and carries no causal interpretation. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.

E Reverse predictive regressions

Table S7 reports subsequent changes in the curve slope on lagged warrant changes and levels. The estimates are imprecise. An inability to reject a zero coefficient is not an equivalence test, a power assessment, or an exclusion of reverse feedback.

Table S7 Reverse predictive associations

Predictor	Coefficient	SE	Webb p	N
Lagged ΔW	0.00008040	0.00004489	0.121	2,504
Lagged $W/100,000$	-0.02607	0.02821	0.416	2,505

Notes. Outcome: next-session change in the fitted curve slope, CNY/tonne/month. The first coefficient is per tonne of lagged net warrant change; the second is per 100,000 tonnes of lagged warrant holdings. Imprecise estimates neither exclude feedback nor establish one-way causality. Standard errors are CRV1 estimates clustered by calendar year. Two-sided bootstrap p-values impose the null and use Webb six-point weights, 9,999 draws, and fixed seed 20260929.